

## Durham Research Online

---

### Deposited in DRO:

06 June 2014

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Bielby, R. and Hill, M.D. and Shanks, T. and Crighton, N.H.M. and Infante, L. and Bornancini, C.G. and Francke, H. and Héraudeau, P. and Lambas, D.G. and Metcalfe, N. and Minniti, D. and Padilla, N. and Theuns, Tom and Tummuangpak, P. and Weilbacher, P. (2013) 'The VLT LBG Redshift Survey - III. The clustering and dynamics of Lyman-break galaxies at  $z \approx 3$ .', *Monthly notices of the Royal Astronomical Society.*, 430 (1). pp. 425-449.

### Further information on publisher's website:

<http://dx.doi.org/10.1093/mnras/sts639>

### Publisher's copyright statement:

This article has been accepted for publication in *Monthly Notices of the Royal Astronomical Society*. ©: 2013 The Authors Published by Oxford University Press on behalf of the Royal Astronomical Society.

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

# The VLT LBG Redshift Survey – III. The clustering and dynamics of Lyman-break galaxies at $z \sim 3^*$

R. Bielby,<sup>1†</sup> M. D. Hill,<sup>1</sup> T. Shanks,<sup>1</sup> N. H. M. Crichton,<sup>1,2</sup> L. Infante,<sup>3</sup>  
C. G. Bornancini,<sup>4</sup> H. Francke,<sup>3</sup> P. Héraudeau,<sup>5,6</sup> D. G. Lambas,<sup>4</sup> N. Metcalfe,<sup>1</sup>  
D. Minniti,<sup>3,7,8</sup> N. Padilla,<sup>3</sup> T. Theuns,<sup>9,10</sup> P. Tummuangpak<sup>1</sup> and P. Weilbacher<sup>11</sup>

<sup>1</sup>Durham University, South Road, Durham DH1 3LE, UK

<sup>2</sup>Max Planck Institute for Astronomy, Knigstuhl 17, D-69117 Heidelberg, Germany

<sup>3</sup>Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile

<sup>4</sup>Instituto de Astronomía Teórica y Experimental (CONICET-UNC), Observatorio Astronómico de Córdoba, Laprida 854, X5000BGR Córdoba, Argentina

<sup>5</sup>Argelander Institut für Astronomie, Auf dem Hügel 71, D-53121 Bonn, Germany

<sup>6</sup>Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands

<sup>7</sup>Vatican Observatory, V00120 Vatican City State, Italy

<sup>8</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544-1001, USA

<sup>9</sup>Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

<sup>10</sup>Department of Physics, University of Antwerp, Campus Groenenborger, Groenenborgerlaan 171, B-2020 Antwerp, Belgium

<sup>11</sup>Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

Accepted 2012 December 17. Received 2012 December 17; in original form 2012 April 13

## ABSTRACT

We present a catalogue of 2135 galaxy redshifts from the VLT LBG Redshift Survey (VLRS), a spectroscopic survey of  $z \approx 3$  galaxies in wide fields centred on background quasi-stellar objects. We have used deep optical imaging to select galaxies via the Lyman-break technique. Spectroscopy of the Lyman-break galaxies (LBGs) was then made using the Very Large Telescope (VLT) Visible Multi-Object Spectrograph (VIMOS) instrument, giving a mean redshift of  $z = 2.79$ . We analyse the clustering properties of the VLRS sample and also of the VLRS sample combined with the smaller area Keck-based survey of Steidel et al. From the semiprojected correlation function,  $w_p(\sigma)$ , for the VLRS and combined surveys, we find that the results are well fit with a single power-law model, with clustering scale lengths of  $r_0 = 3.46 \pm 0.41$  and  $3.83 \pm 0.24 h^{-1}$  Mpc, respectively. We note that the corresponding combined  $\xi(r)$  slope is flatter than for local galaxies at  $\gamma = 1.5$ – $1.6$  rather than  $\gamma = 1.8$ . This flat slope is confirmed by the  $z$ -space correlation function,  $\xi(s)$ , and in the range  $10 < s < 100 h^{-1}$  Mpc the VLRS shows an  $\approx 2.5\sigma$  excess over the  $\Lambda$  cold dark matter ( $\Lambda$ CDM) linear prediction. This excess may be consistent with recent evidence for non-Gaussianity in clustering results at  $z \approx 1$ . We then analyse the LBG  $z$ -space distortions using the 2D correlation function,  $\xi(\sigma, \pi)$ , finding for the combined sample a large-scale infall parameter of  $\beta = 0.38 \pm 0.19$  and a velocity dispersion of  $\sqrt{\langle w_z^2 \rangle} = 420^{+140}_{-160} \text{ km s}^{-1}$ . Based on our measured  $\beta$ , we are able to determine the gravitational growth rate, finding a value of  $f(z = 3) = 0.99 \pm 0.50$  (or  $f\sigma_8 = 0.26 \pm 0.13$ ), which is the highest redshift measurement of the growth rate via galaxy clustering and is consistent with  $\Lambda$ CDM. Finally, we constrain the mean halo mass for the LBG population, finding that the VLRS and combined sample suggest mean halo masses of  $\log(M_{\text{DM}}/M_{\odot}) = 11.57 \pm 0.15$  and  $11.73 \pm 0.07$ , respectively.

**Key words:** galaxies: kinematics and dynamics – cosmology: observations – large-scale structure of Universe.

\*Based on data obtained with the NOAO Mayall 4-m Telescope at Kitt Peak National Observatory, USA (programme ID: 06A-0133), the NOAO Blanco 4-m Telescope at Cerro Tololo Inter-American Observatory, Chile (programme IDs: 03B-0162, 04B-0022) and the ESO VLT, Chile (programme IDs: 075.A-0683, 077.A-0612, 079.A-0442).

†E-mail: rmbielby@googlemail.com

## 1 INTRODUCTION

The large-scale structure of matter presents a crucial guide in understanding the nature and evolution of the Universe. In  $\Lambda$  cold dark matter ( $\Lambda$ CDM), structure in the Universe grows hierarchically through gravitational instability (e.g. Mo & White 1996; Jenkins et al. 1998; Springel, Frenk & White 2006) and testing this model requires the measurement of the matter clustering and the growth of structure across cosmic time (e.g. Springel et al. 2005; Orsi et al. 2008; Kim et al. 2009). We are limited however in our ability to trace the structure of mass given that observations suggest that  $\approx 75$  per cent of the mass density of the Universe is in the form of dark matter.

Although large photometric surveys are beginning to map the overall matter density distribution via its lensing signature (e.g. Massey et al. 2007; Hildebrandt et al. 2012), at present the primary tool in the statistical analysis of the distribution of matter in the Universe remains the study of the clustering statistics of selected galaxy populations. A given galaxy population traces the peaks in the matter distribution and hence provides a biased view of the matter density, which nevertheless can be used to follow the overall growth of structure.

At low redshift, magnitude limited galaxy samples have provided significant tools in probing the clustering properties of the galaxy population (e.g. Norberg et al. 2002; Hawkins et al. 2003), whilst at higher redshifts, photometric selections are required to isolate the required redshift range, for example the luminous red galaxy (LRG),  $BzK$ , extremely red object (ERO), distant red galaxy (DRG) selections. At  $z > 2$ , identifying galaxy populations is primarily reliant on the Lyman-break galaxy (LBG; e.g. Madau et al. 1996; Steidel et al. 1996, 1999) and the Ly $\alpha$  emitter (LAE; e.g. Cowie & Hu 1998; Gawiser et al. 2006, 2007; Ouchi et al. 2008) selections.

In particular, the Lyman-break technique has proven highly successful in surveying the  $z > 2$  Universe. The LBGs represent a large population of star-forming galaxies in the high-redshift Universe. In comparison to LAEs, the LBG selections offer the advantage of both a contiguous and broader range of redshifts, whilst the typically brighter apparent magnitudes of LBGs mean that it is possible to obtain much more detailed information on stellar populations for individual objects, and also to measure a range of interstellar absorption features in rest-frame ultraviolet (UV) spectra.

Steidel et al. (2003) presented a large survey of LBGs in the redshift range  $2.5 < z < 3.5$ , identifying  $\approx 800$  such galaxies based on spectroscopic observations using the Keck I telescope. Adelberger et al. (2003) used this sample to measure the autocorrelation of the LBGs for comparison to the cross-correlation between LBGs and gas as traced by the H I and C IV absorption features in quasar sightlines. They fit the autocorrelation function with a simple power law and reported a clustering length for the galaxies of  $r_0 = 3.96 h^{-1}$  Mpc (with a slope of  $\gamma = 1.55$ ).

Adelberger et al. (2005a) continued from the previous work, presenting an analysis of the clustering properties of galaxies selected photometrically with three different methods including the LBG method. Based on both photometric and spectroscopic samples they found a clustering length of  $r_0 = 4.0 h^{-1}$  Mpc and slope of  $\gamma = 1.6$ , consistent with the previous Keck analyses. Comparison to numerical simulations suggested that such clustering properties were consistent with the LBGs residing in dark matter (DM) haloes with average masses of  $10^{11.2-11.8} M_\odot$ , concluding that the typical LBG will have evolved into an elliptical galaxy at  $z = 0$  and will have an early-type stellar population by  $z \sim 1$ . This was however

contradicted by Conroy et al. (2008) and Bielby et al. (2011, hereafter Paper I), both of whom showed that the clustering evolution of the LBG population may be more complicated, but is likely to produce typical  $L^*$  galaxies at  $z \sim 0$ . Interestingly, this is well complemented by the findings of Quadri et al. (2007, 2008) who show that optically faint/ $K$ -band bright galaxies at  $z \sim 2-3.5$  are far more highly clustered than the optically bright LBG population, and hence suggest that it is this optically faint population missed by the LBG selection that evolves into the massive elliptical population at  $z \sim 0$ . Other observations show consistent measurements of the halo masses in which LBGs reside (e.g. Foucaud et al. 2003; Hildebrandt et al. 2009; Savoy et al. 2011; Jose et al. 2012; Trainor & Steidel 2012). Similarly, the complexity of the evolutionary track of LBGs is supported by recent simulations. For example, González et al. (2012) find that LBGs can be successfully simulated as starbursts triggered by minor mergers, with host halo masses of  $\sim 3 \times 10^{11} h^{-1} M_\odot$ . These are marginally preferentially disc-dominated systems at  $z \sim 3$  that evolve into Milky Way mass galaxies with 50:50 bulge-disc-dominated systems.

Taking the galaxy clustering measurements, it is possible to measure the large-scale dynamics of the galaxy population through redshift-space distortions. For instance, da Ángela, Outram & Shanks (2005b) took the Steidel et al. (2003) Keck spectroscopic sample and used the clustering properties of the LBG population to constrain the cosmological density parameter,  $\Omega_m$ , and the bulk motion properties of the large-scale structure at  $z \approx 3$ . By measuring the 2D clustering of the galaxy distribution, they placed constraints on the infall parameter of  $\beta(z=3) = 0.25^{+0.05}_{-0.06}$  and on the mass density of  $\Omega_m(z=0) = 0.55^{+0.45}_{-0.16}$ . However, the small fields of view available from the Steidel et al. (2003) survey meant the authors could not solve for both the bulk motion and the velocity dispersion, which are degenerate, severely limiting the scope of the results. Paper I improved on these results by combining the data with first galaxy sample from the VLT LBG Redshift Survey (VLRS). By adding  $\approx 1000$  galaxies to the  $z \approx 3$  sample of Steidel et al. (2003) data across much larger fields, they measured the clustering and dynamics of the  $z \approx 3$  LBG population. With the wider fields available, Paper I was able to begin to probe both the small-scale peculiar velocity field and the large-scale bulk motion field. The authors showed that the redshift-space distortions of the  $z \sim 3$  galaxy population are well fit by a model with an infall parameter of  $\beta = 0.48 \pm 0.17$ , which they went on to show is consistent with the standard  $\Lambda$ CDM cosmology. This was similar to a number of other works performed based on redshift distortions at lower redshifts, for example Tegmark et al. (2006), Ross et al. (2007), Guzzo et al. (2008) and Song & Percival (2009), where constraints have been placed on the growth of structure. However, few constraints on this important cosmological measure are available at redshift of  $z \gtrsim 1$ .

In this paper, we add to the previous results of the VLRS presented in Paper I, Crighton et al. (2011, hereafter Paper II) and Shanks, Bielby & Infante (2011). We present new spectroscopic LBG data obtained using the Very Large Telescope (VLT) Visible Multi-Object Spectrograph (VIMOS) instrument, more than doubling both the area covered and the number of spectroscopically confirmed galaxies in the survey. We use the updated survey to measure the clustering and dynamical properties of the  $z \approx 3$  LBG population. Throughout this paper, we use a cosmology given by  $H_0 = 70 \text{ km s}^{-1}$ ,  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $\sigma_8 = 0.8$ . In addition distances are quoted in comoving coordinates in units of  $h^{-1}$  Mpc unless otherwise stated.

## 2 OBSERVATIONS

### 2.1 Survey overview

In order to facilitate an investigation of how  $z \approx 3$  galaxies interact with gas in the intergalactic medium (IGM), the survey comprises observations of several target fields centred on bright  $z > 3$  quasars, since features in the quasi-stellar object (QSO) spectra can provide information on the local IGM. Paper I presented the first five fields of the survey, centred on the following quasars: Q0042–2627 ( $z = 3.29$ ), J0124+0044 ( $z = 3.84$ ), HE 0940–1050 ( $z = 3.05$ ), J1201+0116 ( $z = 3.23$ ) and PKS 2126–158 ( $z = 3.28$ ), hereafter referred to by only the right ascension component of these names. A spectroscopic survey of each of these quasar fields was carried out with the VIMOS on the European Southern Observatory's VLT in Chile (during the ESO periods 75–79). Each field consisted of four subfields (individual pointings with the VLT spectrograph), except for HE 0940 where only three subfields were available at the time of their publication. A VIMOS pointing has a field of view (FoV) of  $16 \times 18 \text{ arcmin}^2$  (see Section 2.4.1), therefore, each quasar field covered  $\approx 32 \times 36 \text{ arcmin}^2$ , or  $\approx 0.32 \text{ deg}^2$ , except for HE 0940 which with three subfields covered  $\approx 0.24 \text{ deg}^2$ .

Building on this initial data set, we present the continuation of these observations since incorporating ESO periods 81 and 82. We have added a further six subfields to HE 0940, tripling its previous area, as well as observations of four new fields, around the quasars Q2359+0653 ( $z = 3.23$ ), Q0301–0035 ( $z = 3.23$ ), Q2231+0015 ( $z = 3.02$ ) and Q2348–011 ( $z = 3.02$ ), with four, four, three and nine subfields, respectively. Table 1 summarizes all the fields of the survey. This includes those presented by Paper I, covering  $1.52 \text{ deg}^2$ , and those presented here, which take the total observed area to  $3.6 \text{ deg}^2$ , more than doubling the previous size.

### 2.2 Imaging

#### 2.2.1 Observations and data reduction

The selection of  $z \approx 3$  LBG candidates was performed using photometry from optical broad-band imaging. The imaging data for

**Table 1.** A summary of the fields making up our  $z \approx 3$  LBG survey. The table gives the name, coordinates and redshift of the QSO on which the fields are roughly centred, as well as the number of subfields (individual VLT VIMOS pointings) with spectroscopic data. The first block of fields was presented by Paper I, the second block is presented in this paper.

Field	RA <sup>a</sup>	Dec. <sup>a</sup>	$z^b$	Subfields	Reference
Q0042–2627	00:44:33.9	–26:11:21	3.29	4	Paper I
J0124+0044	01:24:03.8	+00:44:33	3.84	4	Paper I
HE 0940–1050	09:42:53.4	–11:04:25	3.05	3	Paper I
J1201+0116	12:01:44.4	+01:16:12	3.23	4	Paper I
PKS 2126–158	21:29:12.2	–15:38:41	3.28	4	Paper I
				19	
Q2359+0653	00:01:40.6	+07:09:54	3.23	4	This work
Q0301–0035	03:03:41.0	–00:23:22	3.23	4	This work
Q2231+0015	22:34:09.0	+00:00:02	3.02	3	This work
HE 0940–1050	09:42:53.4	–11:04:25	3.05	6	This work
Q2348–011	23:50:57.9	–00:52:10	3.02	9	This work
				26	

<sup>a</sup>J2000 coordinates of QSO; not necessarily the exact centre of the observed field.

<sup>b</sup>Redshift of the central quasar.

Q2359 and Q0301 were acquired with the Mosaic wide-field imager on the 4-m Mayall telescope at Kitt Peak National Observatory (KPNO) in 2005 September. The Q2231 data are from the Wide Field Camera on the 2.5-m Isaac Newton Telescope (INT) on La Palma, and were observed in 2005 August. All of these observations were carried out in the  $U$ ,  $B$  and  $R$  bands.

The Mosaic imager at KPNO consists of eight  $2k \times 4k$  CCDs arranged into an  $8k \times 8k$  square. With a plate scale of  $0.26 \text{ arcsec pixel}^{-1}$ , this gives a FoV of  $36 \times 36 \text{ arcmin}^2$ . There are 0.5–0.7 mm gaps between the chips, corresponding to gaps of 9–13 arcsec on-sky, so a dithering pattern was used during the observations to provide complete field coverage.  $U$ , Harris  $B$  and Harris  $R$  filters were used.

The Mosaic data were reduced using the MSCRED package in IRAF. The reduction process is described by Paper I, however, we briefly outline the procedure here. Initially a master bias frame is produced for each night's observing. The dome flats and sky flats were then processed using the CCDPROC and MCSPPIL routines, subtracting the bias and eliminating the faint 2600-pixel pupil image artefact. The object frames were processed similarly, subtracting the bias and pupil image and then were flat-fielded using the dome and sky flats. Bad pixels and cosmic rays were masked out of the science frames using the CRREJECT, CRPLUSBPMASK and FIXPIX procedures. Finally, the SWARP software package (Bertin et al. 2002) was used to resample and co-add the frames, producing a final science image.

The HE 0940 and Q2348 data were acquired with the MegaCam imager on the 3.6-m Canada–France–Hawaii Telescope (CFHT). HE 0940 was observed using the CFHT  $u^*$ ,  $g'$ ,  $r'$ ,  $i'$  and  $z'$  bands in 2004 April as part of the observing run 2004AF02 (PI: P. Petitjean), whilst Q2348 was observed in the  $u^*$ ,  $g'$ ,  $r'$  and  $z'$  bands over the period 2004 August–December as part of the observing run 2004BF03 (PI: P. Petitjean). Table 2 gives full details of all the imaging data. For this work we used pre-reduced individual exposures provided by the ELIXIR system at the CFHT Science Archive, which we then stacked using the SCAMP (Bertin 2006) and SWARP (Bertin et al. 2002) software packages.

The Wide Field Camera (WFCam) on the INT comprises four  $2k \times 4k$  CCDs. These are arranged into a  $6k \times 6k$  block with a  $2k \times 2k$  square missing. With  $\approx 1 \text{ arcmin}$  gaps between chips and a pixel scale of  $0.33 \text{ arcsec pixel}^{-1}$ , WFCam has a total FoV of  $\approx 34 \times 34 \text{ arcmin}^2$  ( $0.32 \text{ deg}^2$ ); however, accounting for the incomplete coverage of the field, the total observing area is reduced to  $0.28 \text{ deg}^2$ .

The WFCam observations of Q2231 were made using the RGO  $U$ , Harris  $B$  and Harris  $R$  filters. The RGO  $U$  filter has a central wavelength of  $3581 \text{ Å}$  and a full width at half-maximum (FWHM) of  $638 \text{ Å}$ , making it very similar to the  $U$ -band filter used at KPNO (centre  $3552 \text{ Å}$ , FWHM  $631 \text{ Å}$ ). The  $B$ - and  $R$ -band filters were the same as at KPNO. Therefore, given that the filters are so similar, we will use the same  $UBR$  selection criteria when identifying LBG candidates in either the Mosaic or WFCam data sets.

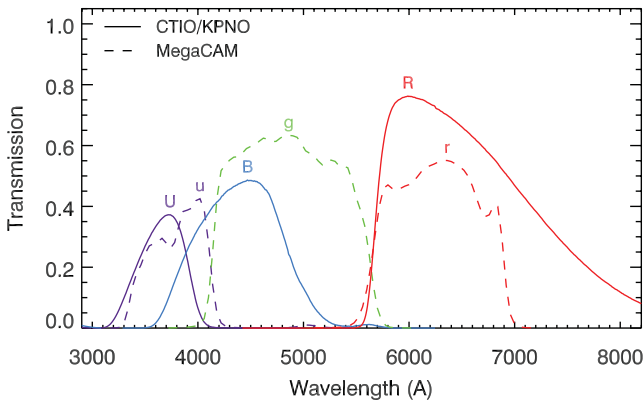
Initial data reduction, including bias removal, flat-fielding and photometric calibration, was performed by the Cambridge Astronomical Survey Unit (CASU). Astrometry calibration and exposure stacking was performed using the SCAMP and SWARP packages.

#### 2.2.2 Filters

As described above, our observations incorporate two different filter combinations. We show both the MegaCam  $u^*g'r'$  and CTIO/KPNO

**Table 2.** Details of imaging observations for the LBG target fields presented in this paper.

Field	RA (J2000)	Dec.	Instrument	Band	Exposure (ks)	Seeing (arcsec)	Completeness (50 per cent Ext/PS)	Dates
Q2359	00:01:44.85	+07:11:56.0	Mosaic (KPNO)	<i>U</i>	19.2	1.46	24.76/25.18	2005 September 29–30
				<i>B</i>	7.2	1.45	25.28/25.73	
				<i>R</i>	6.0	1.15	24.74/25.20	
Q0301	03:03:45.27	−00:21:34.2	Mosaic (KPNO)	<i>U</i>	19.2	1.34	24.93/25.34	2005 September 29–30
				<i>B</i>	6.4	1.28	25.51/26.04	
				<i>R</i>	4.8	1.19	24.59/25.17	
Q2231	22:34:28.00	+00:00:02.0	WFCam (INT)	<i>U</i>	54.0	1.23	25.08/25.52	2005 August 30
				<i>B</i>	13.2	1.01	25.88/26.12	
				<i>R</i>	19.2	1.01	24.75/25.24	
HE 0940	09:42:53.06	−11:02:56.9	MegaCam (CFHT)	<i>u</i> *	6.8	0.99	25.39/25.93	2004 April 14, 21–27
				<i>g</i> '	3.1	0.86	25.54/26.05	
				<i>r</i> '	3.7	0.85	25.08/25.65	
Q2348	23:50:57.90	−00:52:09.9	MegaCam (CFHT)	<i>u</i> *	9.9	0.78	25.97/26.62	August 19–20
				<i>g</i> '	5.5	0.79	25.71/26.29	November 7–10
				<i>r</i> '	4.4	0.75	25.22/25.80	2004 December 15

**Figure 1.** The transmission profiles for the filter combinations used at KPNO/CTIO (solid curves – *UBR*) and CFHT (dashed curves – *ugr*).

*UBR* filter profiles in Fig. 1. The MegaCAM filters have central wavelengths of 3740, 4870 and 6250 Å for the *u*\*, *g*' and *r*' filters, respectively, whilst the CTIO/KPNO filters have central wavelengths of 3570, 4360 and 6440 Å for the *U*, *B* and *R* filters, respectively. These are both well suited to isolating the Lyman break in  $z \sim 3$  galaxies, however, the MegaCAM *u*\* and *g*' filters are marginally redder than the Johnson–Cousins *U* and *B* filters.

Conversions from the MegaCAM filter set to the Sloan Digital Sky Survey (SDSS) filter set are given in the CFHT MegaCAM technical documentation, whilst conversions from the SDSS filter set to the Johnson–Cousins system are given by Fukugita et al. (1996). Combining these two sets of relations gives the following conversions between the two filter systems used in this work:

$$(u^* - g') = 1.05(U - B) + 1.10, \quad (1)$$

$$(g' - r') = 0.57(B - R) - 0.22. \quad (2)$$

These relations are used throughout this paper where comparing the MegaCAM and Johnson–Cousins colours.

### 2.2.3 Photometry

Photometric zero-points for the imaging fields were determined from standard star observations carried out as part of each of the

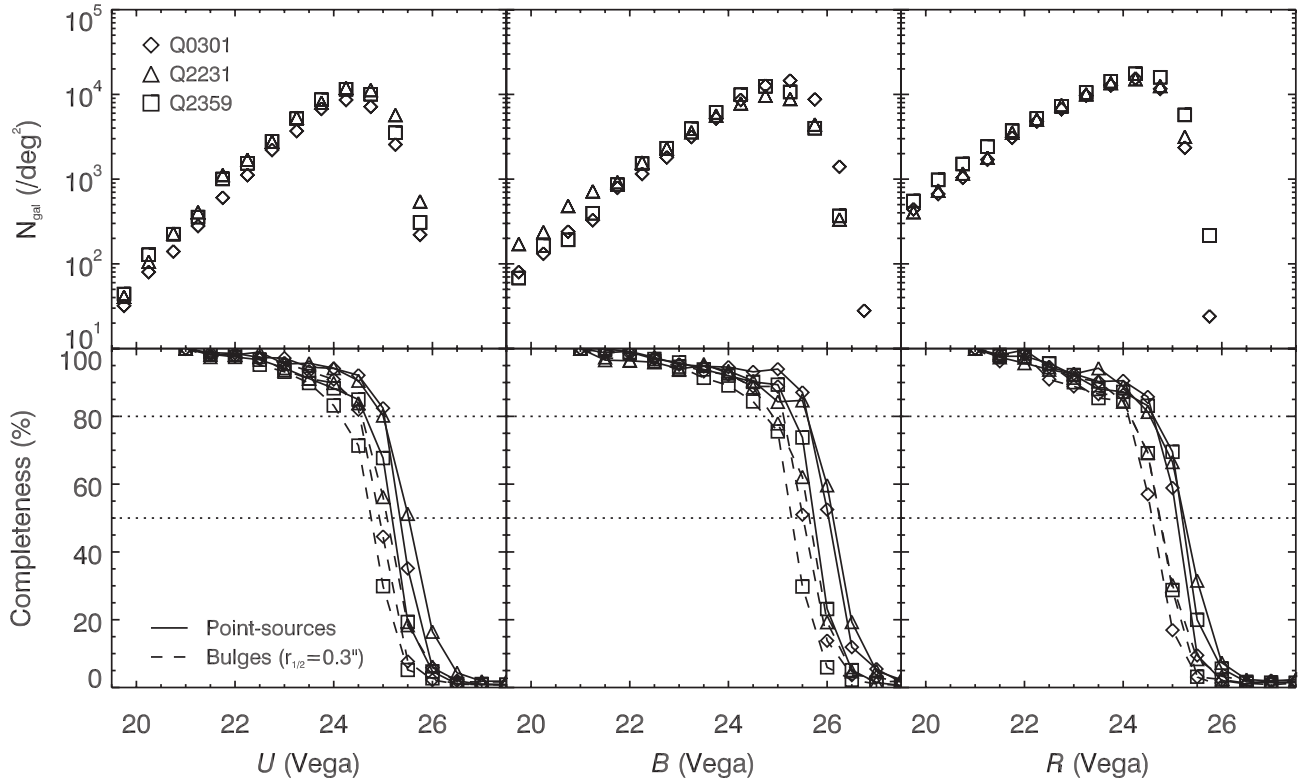
imaging runs. The standard star fields were reduced in the same way as the science frames to ensure consistency. Source detection in the science images was performed with SExtractor (Bertin & Arnouts 1996), using a  $1.5\sigma$  detection threshold and a 5-pixel minimum size.

The  $U_{\text{Vega}}$ ,  $B_{\text{Vega}}$  and  $R_{\text{Vega}}$  band galaxy number counts in the Q0301 (diamonds), Q2231 (triangles) and Q2359 (squares) LBG fields are shown in the top panels of Fig. 2. Stars were removed from these counts at magnitudes brighter than  $\approx 22$  using a limit on the measured half-light radius of the sources. At fainter magnitudes, no attempt to remove stars from the counts was made, as the smallest extended sources become unresolved at the point spread function (PSF) of our fields at such magnitudes. We also show completeness estimates for each image in each field. These are estimated by placing simulated sources at random positions in a given image and measuring the fraction that is successfully extracted with SExtractor (using the same extraction parameters as used to create the full catalogues). In each case we estimate the completeness using both simulated point sources and extended sources, where the extended sources are modelled by a de Vaucouleurs  $r^{1/4}$  profile with a half-light radius of  $r_{1/2} = 0.3$  arcsec. In both cases, the simulated source is convolved with the image PSF before being added to the observation.

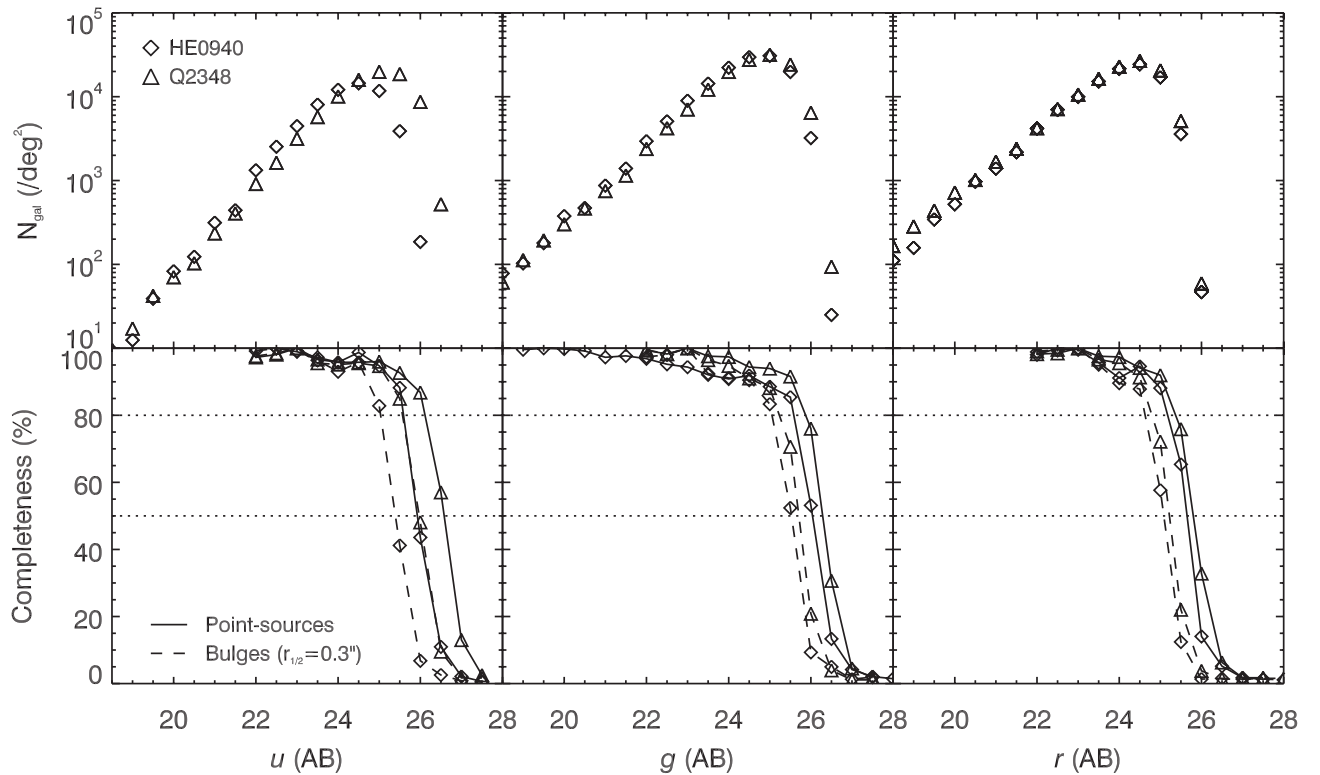
The results of the completeness estimates for the  $UBR_{\text{Vega}}$  filter fields are shown in the lower panels of Fig. 2. The same symbols as the top panels are used for the different fields, whilst the dashed curves show the completeness estimates based on the extended sources and the solid curves show the completeness for the simulated point sources. The 50 per cent limits completion estimates (equivalent to  $\approx 3\sigma$  detection limits) are given in Table 2. Comparing the completeness measurements across the fields, the measurements are relatively consistent with the imaging in each field reaching comparable depths. We note that given the compact nature of the LBG targets, the point source completeness levels should be a good representation of the true completeness. As such all our fields reach depths of  $R > 25$ .

We show the galaxy number counts (top panels) and completeness estimates (lower panels) for the MegaCAM fields in Fig. 3. Again the symbols are consistent between top and lower panels with the diamonds showing the results for the HE 0940 field and the triangles showing the Q2348 field. As before, the solid lines in the lower panels show the completeness estimates for the point-like sources





**Figure 2.** Upper panel: galaxy number counts in the  $U$  (left),  $B$  (middle) and  $R$  (right) band imaging for the fields Q0301 (diamonds), Q2231 (triangles) and Q2359 (squares). Lower panel: estimated completeness for each of the above bands based on simulated point sources (solid lines) and Vaucouleurs profile sources (dashed lines).



**Figure 3.** Number counts (top) and completeness estimates (bottom) in the  $u$  (left),  $g$  (middle) and  $r$  (right) bands from the MegaCAM imaging on the HE 0940–1050 and Q2348–011 fields.

and the dashed lines show the same for the extended sources (which use the same de Vaucouleurs profile as used for the  $UBR_{\text{Vega}}$  fields). Comparing the two fields to each other, the depths reached are comparable in each band, although the HE 0940 is marginally less deep in the  $u$  band by  $\approx 0.5$  mag.

## 2.3 Candidate selection

### 2.3.1 $UBR$ selection

In the Q2359, Q0301 and Q2231 fields, we selected LBG candidates based on their  $U$ ,  $B$  and  $R$  photometry. The criteria used were the same as those used by Paper I, which are based on those of Steidel et al. (2003). There are four groups to the selection, designated LBG\_PRI1, LBG\_PRI2, LBG\_PRI3 and LBG\_DROP and defined as follows:

LBG\_PRI1

- (i)  $23 < R \leq 25$ ;
- (ii)  $0.5 < (U - B) < 4.0$ ;
- (iii)  $(B - R) < 0.8(U - B) + 0.6$ ;
- (iv)  $(B - R) < 2.2$ .

LBG\_PRI2

- (i)  $23 < R \leq 25$ ;
- (ii)  $(U - B) > 0.0$ ;
- (iii)  $(B - R) < 0.8(U - B) + 0.8$ ;
- (iv)  $-1 < (B - R) < 2.7$ ;
- (v)  $\notin \text{LBG\_PRI1}$ .

LBG\_PRI3

- (i)  $23 < R \leq 25$ ;
- (ii)  $-0.5 < (U - B) < 0.0$ ;
- (iii)  $-1.0 < (B - R) < 0.8(U - B) + 0.6$ ;
- (iv)  $\notin \{\text{LBG\_PRI1}, \text{LBG\_PRI2}\}$ .

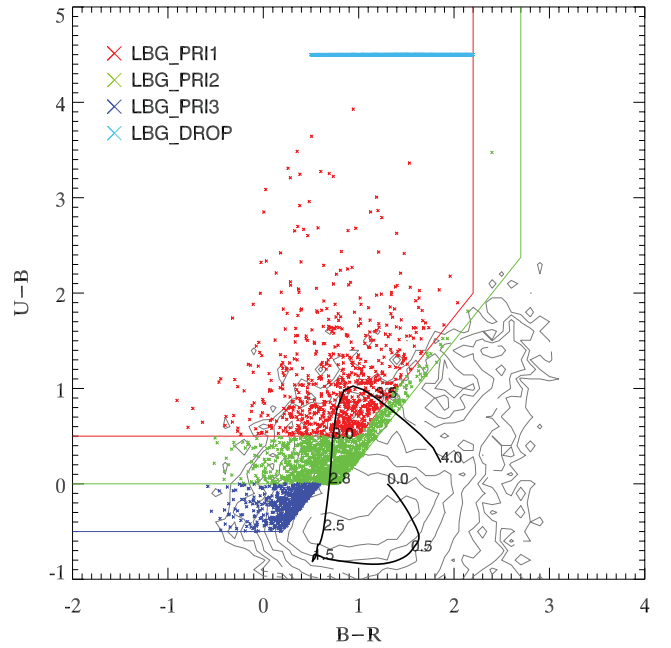
LBG\_DROP

- (i)  $23 < R \leq 25$ ;
- (ii)  $0.5 < (B - R) < 2.2$ ;
- (iii) No detection in  $U$ .

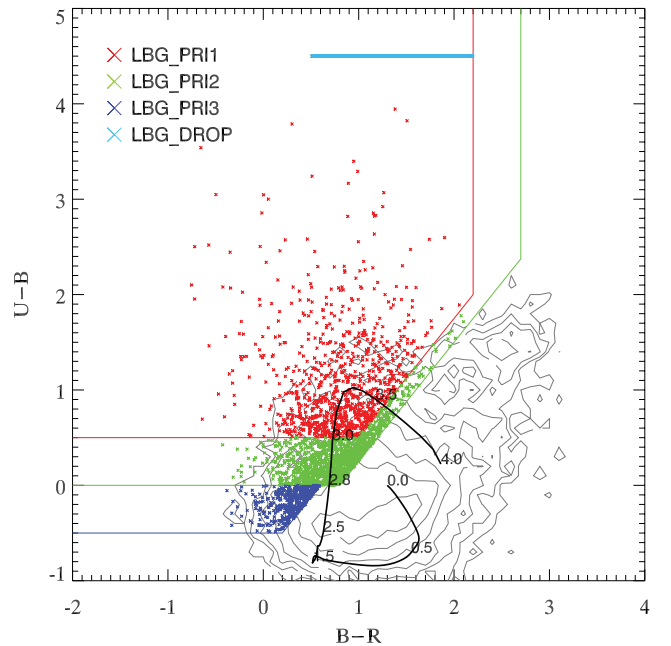
The first three groups represent an order of priority – that is, LBG\_PRI1 candidates are considered more likely to be  $z \approx 3$  LBGs than e.g. LBG\_PRI3 candidates. This is because whereas LBG\_PRI1 tends to select outliers in the  $UBR$  colour–colour plot, the lower priority groups select objects increasingly close to the colour region populated by stars and lower redshift galaxies, and therefore suffer from increased contamination from lower redshift interlopers.

The fourth group is somewhat separate, being for galaxies which are not detected in the  $U$  band. Such sources may be excellent LBG candidates, since it may be that the presence of the Lyman limit in the  $U$  band has made the galaxy extremely faint in this band, such that it ‘drops out’ below the magnitude limit. However, the LBG\_DROP population is also likely to suffer from contamination, in this case because objects with no counterpart in one of the three bands have a higher chance of being spurious sources.

Figs 4–6 show  $UBR$  colour–colour plots for Q2359, Q0301 and Q2231, respectively. In each plot, the LBG\_PRI1, LBG\_PRI2, LBG\_PRI3 and LBG\_DROP candidates are indicated. A model colour–redshift track is also plotted, showing the expected evolutionary path of a star-forming galaxy from  $z = 0$  to 4. This was produced using the Bruzual & Charlot (2003) model, assuming a Chabrier IMF and an exponential star formation rate (SFR) with e-folding time  $\tau = 9$  Gyr. The model indicates that our selection criteria (across all the



**Figure 4.**  $UBR$  colour–colour plot showing candidate selection in Q2359. Objects selected as LBG\_PRI1, LBG\_PRI2, LBG\_PRI3 and LBG\_DROP candidates are shown in different colours as indicated by the legend. The LBG\_DROP candidates have been placed at  $U - B = 4.5$ . The contours show the colour distribution of the rest of the objects in the field.



**Figure 5.** As for Fig. 4, but for Q0301.

priority groups) are predicted to isolate galaxies in the range  $\approx 2.5 < z < 3.8$ . It also suggests that, of the sources that are confirmed as high-redshift LBGs, the LBG\_PRI3s should typically be at a lower redshift than the LBG\_PRI2s, which in turn should be at lower redshift than the LBG\_PRI1s. Paper I noted that this trend was detected in their LBG sample.

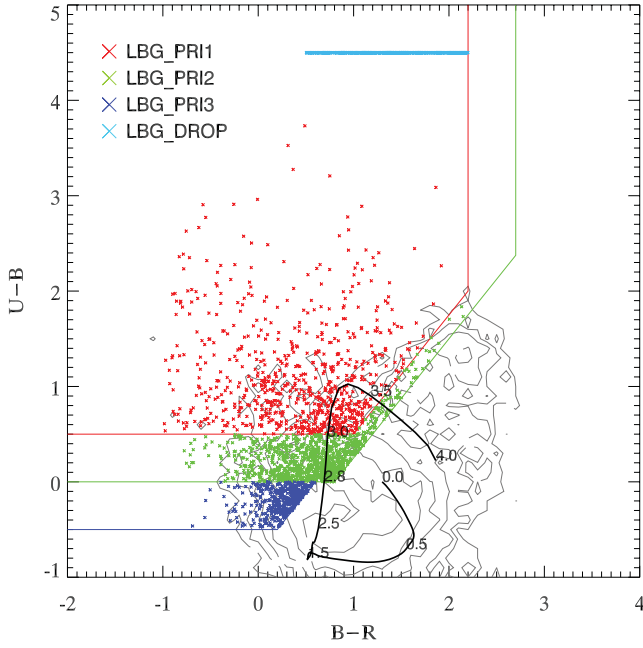


Figure 6. As for Fig. 4, but for Q2231.

### 2.3.2 *ugr* selection

In HE 0940 and Q2348, LBG candidates were selected based on *ugr* photometry. We have therefore adapted the criteria outlined above to account for the different colour bands. Again candidates were selected as either LBG\_PRI1, LBG\_PRI2, LBG\_PRI3 or LBG\_DROP, defined as follows:

LBG\_PRI1

- (i)  $23 < r \leq 25$ ;
- (ii)  $1.4 < (u - g) < 4.0$ ;
- (iii)  $-0.36 < (g - r) < 0.96$ ;
- (iv)  $(g - r) < (u - g) - 1.88$ .

LBG\_PRI2

- (i)  $23 < r \leq 25$ ;
- (ii)  $(u - g) > 1.0$ ;
- (iii)  $-0.36 < (g - r) < 0.96$ ;
- (iv)  $(g - r) < (u - g) - 1.44$ ;
- (v)  $\notin \text{LBG\_PRI1}$ .

LBG\_PRI3

- (i)  $23 < r \leq 25$ ;
- (ii)  $(g - r) < (u - g) - 0.7$ ;
- (iii)  $-0.45 < (g - r) < 0.2(u - g) + 0.1$ ;
- (iv)  $\notin \{\text{LBG\_PRI1}, \text{LBG\_PRI2}\}$ .

LBG\_DROP

- (i)  $23 < r \leq 25$ ;
- (ii)  $-0.36 < (g - r) < 0.96$ ;
- (iii) No detection in *u*.

Figs 7 and 8 show the resulting candidates on *ugr* colour-colour plots, and Table 3 gives the numbers and sky densities of candidates in all five LBG fields.

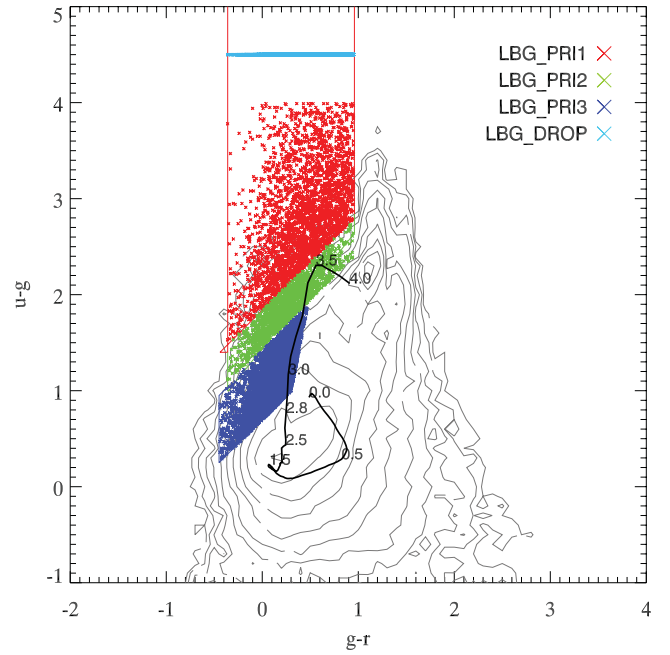


Figure 7. A *ugr* colour-colour plot showing the selection of candidates in HE 0940. Symbols are as in Fig. 4.

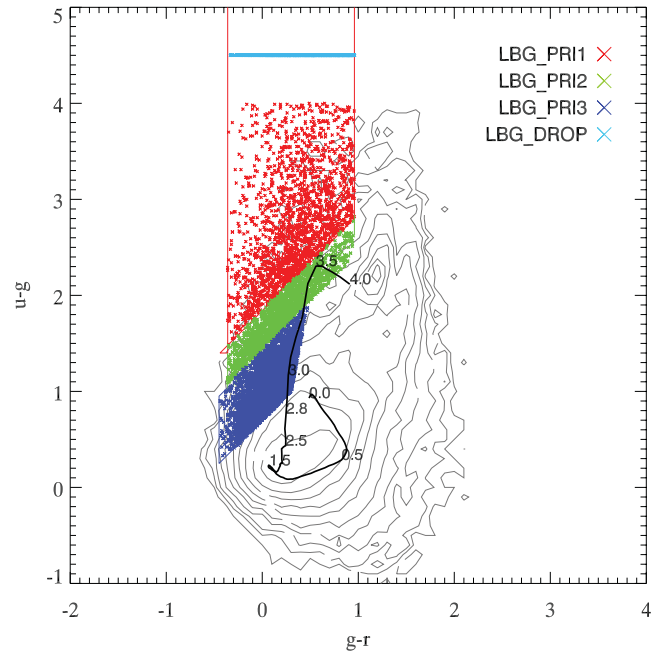


Figure 8. As for Fig. 7, but for Q2348.

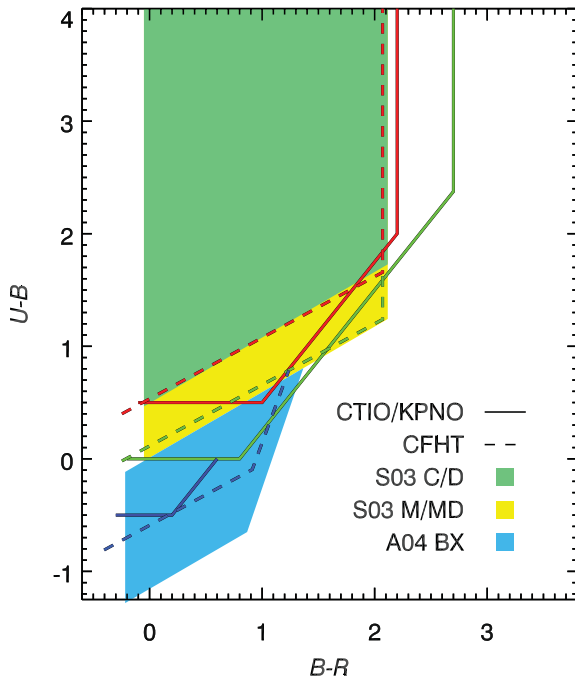
### 2.3.3 Comparing the LBG selections

The *UBR* and *ugr* selections have been developed to mimic the LBG selection of Steidel et al. (2003) and the BX selection of Adelberger et al. (2004). However, given the different sets of filters, the selection functions used here may not perfectly reproduce the original selections. For reference we show the selection functions used here overlaid on the original LBG and BX selections (transformed to the Vega *UBR* system using the relations given by Steidel & Hamilton 1993) in Fig. 9. The CTIO/KPNO filter and CFHT MegaCAM filter (transformed to *UBR* using the relations given on the CFHT



**Table 3.** Numbers and sky densities of  $z \approx 3$  LBG candidates in each of the fields. For each priority class, the first column shows the total number of candidates selected and the second, italicized column shows the density in  $\text{arcmin}^{-2}$ . The last two columns show the figures for all  $z \approx 3$  LBG candidates.

Field	LBG_PRI1		LBG_PRI2		LBG_PRI3		LBG_DROP		All	
Q2359+0653	795	<i>0.61</i>	1130	<i>0.87</i>	549	<i>0.42</i>	709	<i>0.55</i>	3183	<i>2.45</i>
Q0301+0035	891	<i>0.69</i>	1227	<i>0.95</i>	433	<i>0.33</i>	1014	<i>0.78</i>	3565	<i>2.75</i>
Q2231+0015	748	<i>0.65</i>	948	<i>0.82</i>	424	<i>0.37</i>	514	<i>0.45</i>	2634	<i>2.29</i>
HE 0940+1050	2,657	<i>0.69</i>	1896	<i>0.49</i>	5370	<i>1.40</i>	3663	<i>0.96</i>	13 586	<i>3.54</i>
Q2348+011	1843	<i>0.52</i>	1624	<i>0.46</i>	4808	<i>1.35</i>	1850	<i>0.52</i>	10 125	<i>2.84</i>



**Figure 9.** Comparison of the  $UBR$  and  $ugr$  selections compared to the original Steidel et al. (2003, green and yellow filled regions) LBG and Adelberger et al. (2004, cyan filled region) BX  $U_nGR$  selections. All three are presented in the Vega  $UBR$  system.

website<sup>1</sup> combined with those of Fukugita et al. 1996) selections are seen to agree well with the original Steidel et al. (2003) and Adelberger et al. (2004) selections. We note that we do not cover the entire of the Adelberger et al. (2004) BX region as doing so would bring a greater fraction of  $z \lesssim 2$  galaxies. In addition, our  $UBR$  selection boundaries extend somewhat further in the positive  $B - R$  extent. This region is populated by few galaxies in the  $23 < R \leq 25$  range as evidenced by Figs 4–6, but is included to catch faint galaxies that have been scattered in colour space due to photometric errors.

In terms of the resulting space densities, the LBG\_PRI1, LBG\_PRI2 and LBG\_DROP combined for the five fields give a mean space density of  $2.00 \text{ arcmin}^{-2}$  for  $R \leq 25$ , marginally higher than the combination of the C, D, M and MD LBG selections of Steidel et al. (2003) that give a mean sky density of  $\sim 1.8 \text{ arcmin}^{-2}$  for  $R < 25.5$ . Taking the LBG\_PRI3 candidates, we obtain a mean sky density of  $0.37 \text{ arcmin}^{-2}$  in the  $UBR$  fields and  $1.38 \text{ arcmin}^{-2}$  in the  $ugr$  fields. The LBG\_PRI3 selection is intended to provide additional galaxies at  $2 < z < 2.5$  and overlaps to some extent with the Adelberger et al. (2004) BX

selection (as illustrated in Fig. 9). As expected the number densities here for the LBG\_PRI3 selection are much lower than the BX selection, which obtains numbers of  $\sim 5 \text{ arcmin}^{-2}$  at  $R < 25.5$ , due to only sampling a subset of the BX selection.

## 2.4 Spectroscopy

### 2.4.1 Observations

The LBG candidates were targeted in spectroscopic follow-up observations with the VLT VIMOS spectrograph between 2008 September and 2009 December, with programme IDs 081.A-0418(B) (Q2231), 081.A-0418(D) (Q2359), 081.A-0418(F) (Q0301), 082.A-0494(B) (HE 0940) and 082.A-0494(D) (Q2348). The observations were done during dark time in generally good conditions with a typical seeing of  $\approx 1 \text{ arcsec}$  and an airmass of 1.0–1.3. Table 4 gives details of all the fields observed.

The VIMOS instrument (Le Fèvre et al. 2003) comprises four separate CCDs, each with a FoV of  $7 \times 8 \text{ arcmin}^2$ . These four arms are arranged in a  $2 \times 2$  grid with a  $\approx 2 \text{ arcsec}$  gap between each CCD, giving a total  $16 \times 18 \text{ arcmin}^2$  FoV as quoted previously. Of this  $288 \text{ arcmin}^2$  field,  $224 \text{ arcmin}^2$  is covered by the detector.

Our observations utilized the low-resolution blue (LR\_Blue) grism and the order sorting blue (OS\_Blue) filter, resulting in a wavelength range of  $3700\text{--}6700 \text{ \AA}$ , blazing at  $\approx 4000 \text{ \AA}$ . This wavelength range is ideal for our survey, detecting the Ly $\alpha$  line at  $2.0 < z < 4.5$ . The resolving power of the spectrograph in this configuration is  $R = 180$  assuming a 1-arcsec slit (as used in these observations), which gives a resolution element of  $\Delta\lambda \approx 22 \text{ \AA}$  at the blaze wavelength. The spectral dispersion is  $5.3 \text{ \AA pixel}^{-1}$ .

The slit masks were designed using the VMPS software which is standard for VIMOS observations. The aims for mask design are (a) to maximize the number of observed targets, (b) to favour higher priority targets and (c) to ensure slits are of sufficient size to allow a robust sky subtraction. Since these aims are frequently in conflict with one another, the mask design process is one of attempting to optimize the observations to satisfy all three as much as possible. Point (c) is addressed by setting a minimum slit length of 8 arcsec (40 pixels given the pixel scale of  $0.205 \text{ arcsec pixel}^{-1}$ ). Slit length was increased as much as possible where such an increase would not prevent the observation of an additional target – that is, where it did not conflict with point (a). Finally, in order to optimize slit allocation, some targets were added to fill gaps that fulfilled the given selection criteria, but with fainter  $R$  magnitudes down to a limit of  $R = 25.5$ .

With the LR\_Blue grism, each spectrum spans 640 pixels along the dispersion axis. Assuming a 40 pixels slit width as specified above, this would allow for a possible total of over 300 slits on the full  $4k \times 2k$  detector. This is however not practically achievable given the density of LBG candidates, and is hampered further by

<sup>1</sup> <http://www3.cadc-ccda.hia-ihp.nrc-cnrc.gc.ca/megapipe/docs/filters.html>

**Table 4.** Details of spectroscopic observations for the LBG target fields presented in this paper.

Field	Subfield	RA <sup>a</sup>	Dec. <sup>a</sup>	Exposure <sup>b</sup>	Airmass	Seeing (arcsec)	Dates
Q2359	f1	00:01:09.94	+07:03:26.8	10	1.1–1.2	0.6–1.7	2008 September 23–25
Q2359	f2	00:01:12.92	+07:16:39.2	10	1.1–1.3	0.6–1.3	2008 October 3, 20–21
Q2359	f3	00:02:11.50	+07:15:33.9	10	1.2–1.4	0.6–1.3	2008 November 3, 21–25
Q2359	f4	00:02:12.89	+07:02:19.1	10	1.2–1.3	0.5–1.1	2008 November 26–30
Q2231	f1	22:34:28.19	−00:06:03.3	10	1.1–1.2	0.5–1.0	2008 October 23, 28
Q2231	f2	22:34:28.55	+00:06:13.2	10	1.1–1.2	0.4–0.9	2008 October 21–22
Q2231	f3	22:33:39.51	−00:06:10.8	10	1.1–1.2	0.5–1.0	2008 August 3, July 27
Q0301	f1	03:04:20.12	−00:14:28.8	12	1.1–1.3	0.7–1.2	2008 October 23, 31
Q0301	f2	03:03:10.27	−00:16:18.7	12	1.1–1.2	0.7–1.5	2008 November 21–23
Q0301	f3	03:03:15.41	−00:30:40.0	12	1.1–1.2	0.7–1.5	2008 November 25–26
Q0301	f4	03:04:15.56	−00:28:59.1	12	1.1–1.2	0.7–1.2	2008 September 24, October 1, 7
HE 0940	f4	09:42:10.00	−10:54:30.3	11.2	1.0–1.3	0.8–1.5	2009 February 1
HE 0940	f5	09:43:07.47	−11:24:50.3	11.2	1.0–1.3	0.7–1.4	2009 February 3
HE 0940	f6	09:41:59.99	−11:24:50.4	11.2	1.0–1.2	0.5–1.2	2009 February 20–21
HE 0940	f7	09:44:14.99	−11:24:49.9	11.2	1.0–1.2	0.5–1.3	2009 February 22, 24
HE 0940	f8	09:43:21.49	−10:41:00.5	11.2	1.0–1.2	0.5–1.0	2009 February 26–27
HE 0940	f9	09:42:09.99	−10:40:59.8	11.2	1.0–1.3	0.5–1.2	2009 February 2
Q2348	f1	23:51:50.08	−00:54:21.9	11.5	1.1–1.2	0.5–1.7	2009 July 23–25
Q2348	f2	23:50:45.09	−00:54:22.2	11.5	1.0–1.1	0.5–1.0	2009 July 19–20
Q2348	f3	23:49:40.07	−00:54:22.6	11.5	1.0–1.2	0.4–0.8	2009 July 27
Q2348	f4	23:51:50.12	−00:37:31.6	11.5	1.1–1.2	0.5–1.5	2009 August 20–21
Q2348	f5	23:50:45.05	−00:37:31.5	11.5	1.0–1.2	0.7–1.4	2009 September 16–20
Q2348	f6	23:49:40.00	−00:37:32.0	11.5	1.1–1.2	0.8–1.3	2009 September 24–25
Q2348	f7	23:51:50.12	−01:07:31.4	11.5	1.1–1.2	0.7–1.0	2009 October 12, 20
Q2348	f8	23:50:45.00	−01:07:32.0	11.5	1.1–1.3	0.7–1.3	2009 November 22, December 10
Q2348	f9	23:49:40.00	−01:07:32.0	11.5	1.1–1.3	0.8–1.5	2009 November 15–22

<sup>a</sup>J2000 coordinates of subfield centre.<sup>b</sup>In ks.

the need to select high-priority candidates (point b), which have an even lower sky density. Our final slit masks therefore typically contain some 50–70 slits per quadrant.

#### 2.4.2 Data reduction

The spectroscopic data have all been reduced using the VIMOS ESOREX reduction pipeline. Using bias frames, flat-fields and arc lamp exposures taken for each mask during each observing run, the pipeline generates bias-subtracted, flat-fielded, wavelength-calibrated science frames consisting of a series of 2D spectra. Following Paper I we use the IMCOMBINE procedure in IRAF to combine the reduced frames from each observing block, generating a master science frame for each quadrant of each field. When combining the frames we use the CRREJECT mode, designed to remove cosmic rays by rejecting pixels with significant positive spikes. We have also used the AVSIGCLIP rejection mode with a rejection threshold of  $3\sigma$ , and find that our results are not significantly affected, suggesting that our results do not depend strongly on the parameters used to combine the science frames at this stage.

We extract 1D spectra from the reduced, combined 2D spectra using the IDL routine SPECPLAT. One-dimensional object and sky spectra are found by averaging across the respective apertures, and the sky spectrum is then subtracted from the object spectrum to give a final spectrum for the object.

In some cases there remain significant sources of contamination in the final object spectrum. These can arise from bad pixels, either in the object or sky aperture, from contamination from the zeroth-order from other slits, or more frequently from the bright sky emission lines [O I] 5577 Å, [Na I] 5990 Å and [O I] 6300 Å;

in either case, the resulting contamination may manifest itself as either a positive or a negative spike in the spectrum. Such artefacts are, however, easily spotted during a routine inspection of the 2D spectrum.

#### 2.5 Identification of targets

Every source targeted for spectroscopic observation is inspected visually, in both the 2D and 1D spectra, to determine where possible a redshift and classification. Sources are classified as either  $z \approx 3$  LBGs, low-redshift galaxies, QSOs or Galactic stars. The LBGs are divided into those showing Ly $\alpha$  emission (designated LBe) and those showing Ly $\alpha$  absorption (LBA). QSOs are determined by the presence of typical active galactic nucleus (AGN) emission features, in particular Ly $\alpha$  and C IV. Stars are classified by comparison to template spectra: in particular we check for A, F, G, K and M stars.

In determining the redshift and classification the spectral feature primarily used in the case of LBGs is the Ly $\alpha$  emission/absorption line at 1216 Å; for lower redshift galaxies it is the [O II] emission line at 3727 Å. In addition to these, some of the following features are used.

For  $z \approx 3$  LBGs:

- (i) Lyman limit, 912 Å;
- (ii) Ly $\beta$  emission/absorption, 1026 Å;
- (iii) O VI 1032 Å, 1038 Å;
- (iv) Ly $\alpha$  forest, <1215.67 Å;
- (v) Ly $\alpha$  emission/absorption, 1215.67 Å;
- (vi) interstellar medium (ISM) absorption lines:
  - (a) Si II 1260.4 Å;

- (b) O I + Si II 1303 Å;
- (c) C II 1334 Å;
- (d) Si IV doublet 1393 Å and 1403 Å;
- (e) Si II 1527 Å;
- (f) C IV doublet absorption, 1548–1550 Å;
- (g) Fe II 1608 Å;
- (h) Al II 1670 Å.

For low- $z$  galaxies:

- (i) CN absorption 3833 Å;
- (ii)  $K$ -band absorption 3934 Å;
- (iii) HK break 4000 Å;
- (iv) H $\delta$  emission 4102 Å;
- (v) H $\beta$  emission/absorption 4861 Å;
- (vi) O III emission 4959 Å;
- (vii) O III emission 5007 Å.

The presence of the HK break causes these interlopers to appear fairly frequently in our spectroscopic samples, since these features mimic the Lyman break on which our selection is based. The ISM absorption features listed above for LBGs are therefore of considerable importance in identifying genuine  $z \approx 3$  galaxies. For every target which is identified, we assign a quality parameter to the redshift determined, in the range  $0 \leq Q \leq 1$ . A quality of  $Q \leq 0.4$  indicates that a possible redshift has been determined, but is not considered a robust measurement. Above this, for LBGs, the quality parameters indicate that the redshift is based on the following features.

- (i)  $Q = 0.5$  – a spectral break with some weak Ly $\alpha$  emission/absorption and low signal-to-noise ratio (SNR) ISM absorption features, or strong Ly $\alpha$  emission but with no detected continuum.
- (ii)  $Q = 0.6$  – a spectral break with high-SNR Ly $\alpha$  emission/absorption plus low-SNR ISM absorption features.
- (iii)  $Q = 0.7$  – a spectral break with high-SNR Ly $\alpha$  emission/absorption plus unambiguous, high-SNR ISM absorption features.
- (iv)  $Q = 0.8$  – a spectral break with high-SNR Ly $\alpha$  emission/absorption plus high-SNR absorption and lower SNR emission lines (e.g. Si II 1265 Å, 1309 Å; He II 1640 Å).
- (v)  $Q = 0.9$  – as for  $Q = 0.8$ , but reserved for highest SNR objects only.

## 2.6 LBG sample

### 2.6.1 Sky densities, redshift distributions and completeness

In total, the VLRS now consists of 2135 spectroscopically confirmed LBGs in  $\approx 10\,000$  arcmin<sup>2</sup> (a density of  $\approx 0.20$  arcmin<sup>-2</sup>). 1994 of these are within a magnitude limit of  $R \leq 25$ , whilst the remainder form a non-uniform sample of  $25 < R \leq 25.5$  LBGs that were observed as part of optimizing slit allocations in the spectroscopic observations. We show the sky distribution of LBGs from both Paper I (open grey circles) and this paper (filled black circles) for all nine VLRS fields in Fig. 10. Known QSOs in the fields are also plotted (cyan stars). We present examples of the first six LBGs in each of the fields in Tables 5–9. The full tables will be made available online at <http://star-www.dur.ac.uk/~bielby/vlrs/>. The total numbers of  $R \leq 25$ ,  $Q \geq 0.5$  sources identified in each of the five fields presented here are given in Table 10.

For the  $z > 2$  sample, the VLRS now consists of 944  $Q = 0.5$ , 492  $Q = 0.6$ , 318  $Q = 0.7$ , 147  $Q = 0.8$  and 93  $Q = 0.9$  galaxies at a magnitude limit of  $R \leq 25$ .

Fig. 11 shows the  $n(z)$  distributions of all sources with measured redshifts in each of the five LBG fields. The figure shows that LBGs in HE 0940 and Q2348, where LBGs were selected in *ugr*, have higher average redshifts than in the *UBR*-selected fields, suggesting that the *ugr* criteria bias the selection towards higher  $z$ . It is also notable from Table 10 that the *ugr* selection appears to include more Galactic stars. Future *ugr*-selected LBG surveys may wish to alter our colour criteria to better avoid stellar interlopers.

Fig. 11 also shows  $n(z)$  for the subsets of sources with  $Q \geq 0.6$  and  $\geq 0.7$ . We note that in any given field, the distributions of sources at  $Q \geq 0.5$ ,  $\geq 0.6$  or  $\geq 0.7$  are approximately the same – the LBGs with higher ID qualities are not skewed to lower or higher redshift, for example – suggesting that the redshift distributions shown are fairly robust. The average redshifts and standard deviations are given in Table 11. The redshift distribution of the full LBG sample has a mean redshift of  $2.79 \pm 0.01$  and a standard deviation of 0.34, and is shown in the lower right-hand panel of Fig. 11.

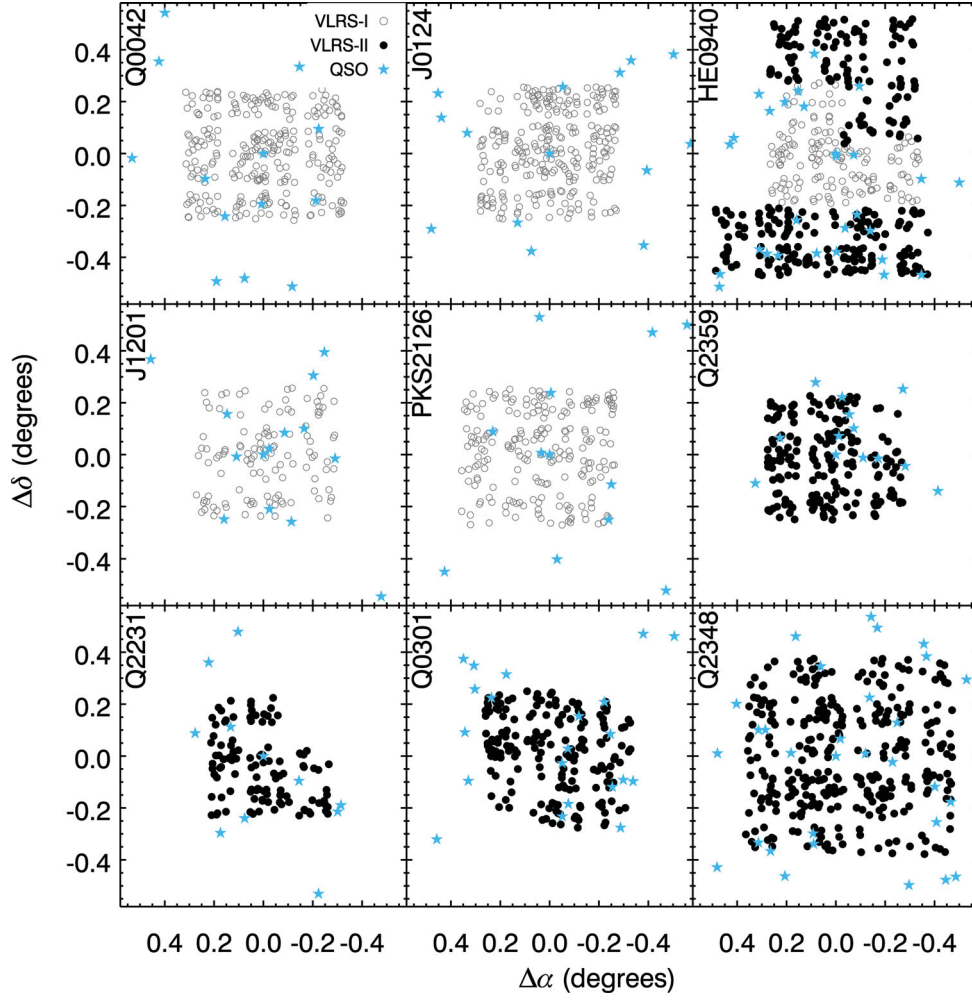
Fig. 12 shows, as anticipated in Section 2.3, that candidates selected as LBG\_PRI1 lie at higher redshift than the LBG\_PRI2 candidates, which are in turn at higher redshift than LBG\_PRI3S. Quantitatively, we find that the LBG\_PRI1S have a mean redshift of  $\bar{z} = 2.83 \pm 0.02$ , the LBG\_PRI2S have  $\bar{z} = 2.71 \pm 0.02$  and the LBG\_PRI3S have  $\bar{z} = 2.61 \pm 0.02$ . The LBG\_DROP candidates are shown in a separate panel for clarity, and have the highest mean redshift of all the groups, with  $\bar{z} = 2.93 \pm 0.02$ .

In total, for the  $R \leq 25$  targets, we make successful identifications in  $\approx 30$  per cent of VIMOS slits and of those identified, 62 per cent are identified as  $z > 2$  galaxies. In terms of the unidentified fraction, these are likely predominantly faint  $z > 2$  galaxies (most likely dominated by dusty, absorbed galaxies with no Ly $\alpha$  emission) and relatively featureless  $1 \lesssim z \lesssim 2$  galaxies. We note that the 38 per cent contamination rate is somewhat higher than that quoted for the Steidel et al. (2003) and subsequent samples. This is in part likely the result of the shallower depths and differing filters used in the colour selections. Additionally, following the results of other authors (e.g. Reddy et al. 2008), it is likely that the faint population that has avoided identification in our observations is less prone to contamination and as such likely has a higher percentage of  $z \sim 3$  galaxies than the 62 per cent measured for the sample in which we could successfully identify spectral features.

Breaking the contamination level into the different selections, we find that the LBG\_PRI1, LBG\_PRI2, LBG\_PRI3 and LBG\_DROP samples have contamination rates of 32.5, 35.6, 38.2 and 40.3 per cent, respectively. Based on these recovered levels of contamination (and making the simplifying assumption that this applies to the faint unidentified spectroscopic sample), gives an average sky density across our fields of  $\approx 1.8$  arcmin<sup>-2</sup> for all samples and  $\approx 1.3$  arcmin<sup>-2</sup> excluding the LBG\_PRI3 sample. Based on the volumes probed and the redshift distribution, these sky densities correspond to sky densities of  $\sim 4.0 h^3 \text{ Mpc}^{-3}$ .

## 2.7 Galaxy redshifts

In star-forming galaxies such as those presented here, the observed Ly $\alpha$  emission is redshifted relative to the intrinsic galaxy redshift, while the interstellar absorption lines are blueshifted (see e.g. Shapley et al. 2003). In Paper I, we used the transformations given by Adelberger et al. (2005b) in order to correct from the redshifts of the UV features to the intrinsic galaxy redshifts. These have now been superseded by those determined in Steidel et al. (2010), which we use in this paper and also apply to our previous data from Paper I.



**Figure 10.** Sky distribution of LBGs and QSOs in the VLRS fields. Grey open circles show LBGs presented in Paper I, black filled circles show LBGs identified in the current work and cyan stars show known QSOs.

**Table 5.** Example LBG identifications in the Q2359+0653.

ID	RA (J2000)	Dec.	$U - B$	$B - R$	$R$	$z_{Ly\alpha}$	$z_{ISM}$	$Q_{ID}$
VLRS J000139.85+070221.66	0.4160563	7.0393505	0.63	0.76	23.5500	2.4762	2.4682	0.5
VLRS J000133.54+070127.57	0.3897395	7.0243263	0.13	0.15	25.3600	2.5707	2.5603	0.5
VLRS J000118.84+070106.55	0.3285175	7.0184855	1.29	1.21	23.8900	3.0374	3.0294	0.5
VLRS J000131.05+070106.56	0.3793770	7.0184898	0.58	0.65	25.3900	2.7910	2.7967	0.5
VLRS J000141.27+070106.35	0.4219468	7.0184293	1.51	0.12	24.3800	2.6508	2.6428	1.0
VLRS J000140.69+070044.11	0.4195270	7.0122533	0.47	0.65	23.7800	2.8162	2.8082	0.5
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...

In Paper I, we estimated the errors on the LBG redshifts using simulated spectra. Here, we extend the investigation into the redshift errors in our survey by using duplicate redshift measurements. The fields presented here, particularly Q2348, were designed with overlapping regions and consequently there are some LBG candidates which were observed in more than one mask. In cases where these duplicated targets are confirmed as LBGs, this provides two independent redshift measurements for the same LBG, and thus a direct observational test of the redshift measurement accuracy.

Fig. 13 shows the  $\Delta z$  distribution for the LBGs with duplicate observations, where  $\Delta z = |z_1 - z_2|$  is the difference between the

two redshift measurements. A total of 20 objects were classified as LBGs in two separate observations; of these, Fig. 13 indicates that 16 had fairly small errors of  $\Delta z < 0.02$  (of which 13 had very small errors of  $\Delta z < 0.005$ ), while four had considerably larger errors. In addition to these 20 objects, we have also searched a region of our Q0301 field which overlaps with Steidel et al. (2003) survey for any LBGs which were identified in both surveys: we find three such objects, and the redshift differences for these galaxies are also indicated in Fig. 13.

The standard deviation of the 20  $\Delta z$  values measured for duplicate observations in our survey is  $\sigma = 0.036$ , corresponding to a velocity



**Table 6.** Example LBG identifications in the Q0301–0035.

ID	RA (J2000)	Dec.	$U - B$	$B - R$	$R$	$z_{\text{Ly}\alpha}$	$z_{\text{ISM}}$	$Q_{\text{ID}}$
VLRS J030434.85–001549.27	46.1452103	–0.2636854	0.59	0.61	24.4700	2.6132	2.6041	0.7
VLRS J030435.40–001607.15	46.1474953	–0.2686527	1.56	1.45	23.4100	2.5969	2.6157	0.7
VLRS J030439.49–001619.35	46.1645317	–0.2720422	1.08	0.62	24.5100	2.9570	2.9490	0.5
VLRS J030438.22–001647.63	46.1592560	–0.2798966	–0.22	0.37	23.8600	2.7098	2.7292	0.6
VLRS J030435.86–001654.14	46.1494179	–0.2817046	0.80	0.92	25.0700	2.8887	2.8807	1.0
VLRS J030426.38–001701.38	46.1099281	–0.2837157	0.21	0.56	24.4900	2.4651	2.4571	0.6
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...

**Table 7.** Example LBG identifications in the HE 0940–1050.

ID	RA (J2000)	Dec.	$u - g$	$g - r$	$r$	$z_{\text{Ly}\alpha}$	$z_{\text{ISM}}$	$Q_{\text{ID}}$
VLRS J094225.83–105744.50	145.6076355	–10.9623623	2.57	0.35	23.6400	2.8804	2.8810	0.5
VLRS J094240.69–105753.44	145.6695251	–10.9648447	1.88	0.35	24.1100	3.1456	3.1376	0.5
VLRS J094220.01–105900.05	145.5833588	–10.9833469	–	–0.14	24.3800	2.2010	2.1930	0.5
VLRS J094217.39–105923.95	145.5724640	–10.9899855	–	0.16	23.8500	2.5153	2.5073	0.5
VLRS J094217.51–105935.92	145.5729675	–10.9933100	–	0.79	24.2600	2.8139	2.8119	0.6
VLRS J094242.29–110121.16	145.6762085	–11.0225439	0.74	–0.03	23.9900	2.4613	2.4595	0.5
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...

**Table 8.** Example LBG identifications in the Q2231+0015.

ID	RA (J2000)	Dec.	$U - B$	$B - R$	$R$	$z_{\text{Ly}\alpha}$	$z_{\text{ISM}}$	$Q_{\text{ID}}$
VLRS J223439.00+000341.29	338.6625061	0.0614693	–	0.72	24.1500	2.8428	2.8291	0.6
VLRS J223459.87+000308.07	338.7494507	0.0522424	1.01	0.78	23.7400	2.4879	2.4789	0.5
VLRS J223450.20+000232.38	338.7091675	0.0423284	–	1.92	23.7800	2.8934	2.8927	0.7
VLRS J223459.03+000051.79	338.7459717	0.0143855	–	0.54	24.3200	2.8037	2.7957	0.9
VLRS J223442.76–000028.59	338.6781616	–0.0079414	0.32	1.01	23.6900	2.1897	2.1817	0.5
VLRS J223447.81–000041.63	338.6991882	–0.0115636	–	0.87	24.8000	2.8874	2.8735	0.5
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...

**Table 9.** Example LBG identifications in the Q2348–011.

ID	RA (J2000)	Dec.	$u - g$	$g - r$	$r$	$z_{\text{Ly}\alpha}$	$z_{\text{ISM}}$	$Q_{\text{ID}}$
VLRS J235206.92–005646.70	358.0288391	–0.9463067	1.30	0.30	24.4600	3.1430	2.3707	0.5
VLRS J235200.98–005903.11	358.0040894	–0.9841969	2.50	0.46	24.2400	3.3532	3.3435	0.5
VLRS J235201.68–010002.18	358.0069885	–1.0006067	1.23	0.31	24.9700	2.8632	2.8603	0.5
VLRS J235155.14–010104.04	357.9797363	–1.0177902	1.66	0.15	23.8600	2.7523	2.7503	0.7
VLRS J235209.28–005535.49	358.0386658	–0.9265237	1.28	0.13	24.0700	2.6274	2.6194	0.7
VLRS J235202.62–004747.89	358.0109253	–0.7966368	2.11	0.32	24.2700	3.0640	3.0646	0.5
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...

error of  $\approx 2800 \text{ km s}^{-1}$  assuming a redshift of  $z = 2.8$ , the sample mean. However, this misrepresents the true error in our redshift measurements, since it is skewed by the four sources with very high  $\Delta z$ . These four values do not represent redshift *measurement* errors, rather catastrophic outliers. In the cases where we find large  $\Delta z$  values, the error does not arise due to uncertainty in the peak wavelength, but in uncertainty over which spectral feature is actually  $\text{Ly}\alpha$ . In these cases, different spectral features have been identified as  $\text{Ly}\alpha$ , leading to large  $\Delta z$ . These are therefore better characterized

as *identification* errors, in that two different solutions have been reached in the two observations.

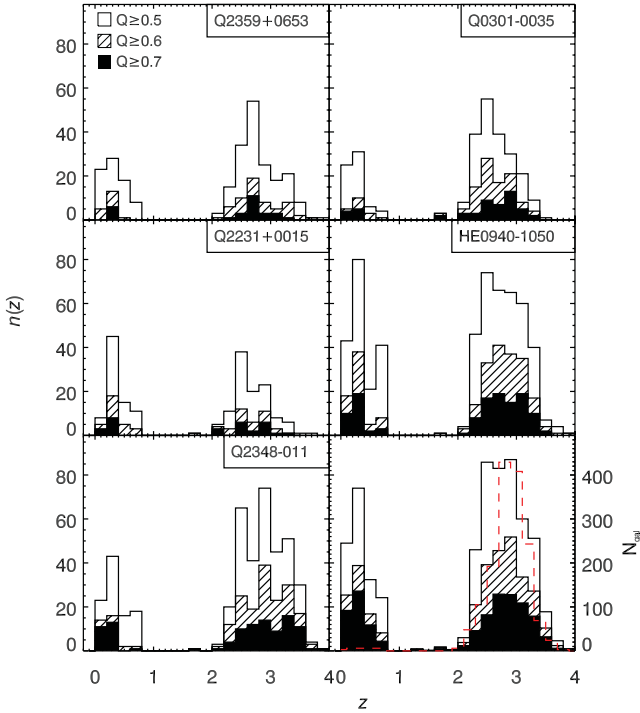
For the 16 duplicated targets shown in Fig. 13(b), the same feature has been identified as  $\text{Ly}\alpha$  and therefore the  $\Delta z$  for these objects gives an indication of the measurement error. The standard deviation for these objects is  $\sigma = 0.005$ , corresponding to  $\Delta v \approx 380 \text{ km s}^{-1}$ .

The suggestion, therefore, is that  $\approx 80$  per cent of our LBGs have redshift measurement errors of  $\Delta v \leq 400 \text{ km s}^{-1}$ , while the other 20 per cent may have larger errors. This problem, however,



**Table 10.** Numbers of objects in each target field spectroscopically identified at  $Q \geq 0.5$  and  $R \leq 25$  as either high- $z$  LBGs,  $z < 2$  galaxies, AGN/QSOs, stars or unidentified.

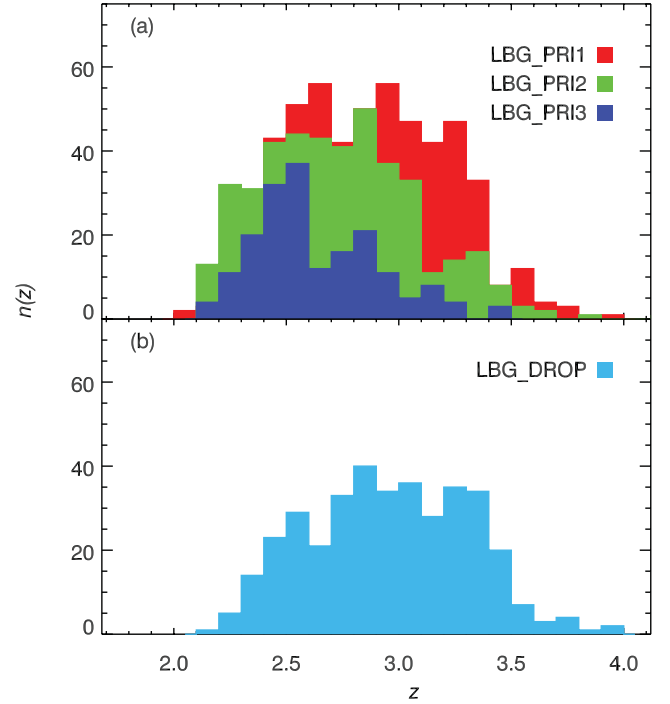
Field	$z \approx 3$ LBGs	$z < 2$ galaxies	QSO/AGN	Stars
Q2359	143 (0.18 arcmin $^{-2}$ )	67	5	8
Q0301	164 (0.21 arcmin $^{-2}$ )	61	10	13
Q2231	108 (0.18 arcmin $^{-2}$ )	80	6	18
HE 0940	358 (0.30 arcmin $^{-2}$ )	186	4	48
Q2348	303 (0.17 arcmin $^{-2}$ )	100	11	34
Total	1076 (0.21 arcmin $^{-2}$ )	494	36	121



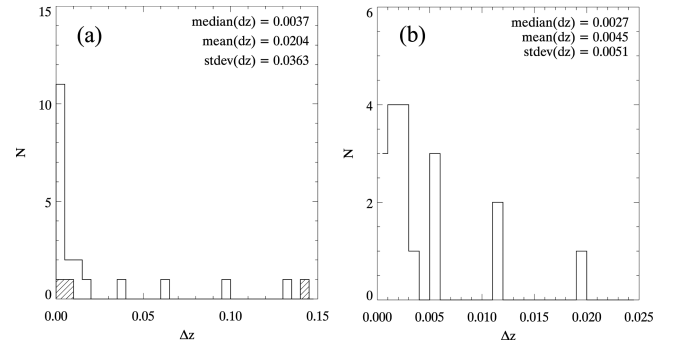
**Figure 11.** Redshift distributions of identified sources in each of the target fields (first five panels) and for the whole sample of 1994  $R \leq 25$  galaxies (bottom right-hand panel). In each panel we show the distribution for the full sample ( $Q \geq 0.5$ ), as well as for the  $Q \geq 0.6$  (hashed) and  $Q \geq 0.7$  (filled) subsets, where  $Q$  is the redshift quality parameter. In the final panel, we also show the redshift distribution of the Steidel et al. (2003) galaxies used in this work (dashed red line).

**Table 11.** Redshift distribution statistics for spectroscopically confirmed,  $R \leq 25$ ,  $Q \geq 0.5$  LBGs in the five observed fields. In each case the mean redshift  $\bar{z}$  (with standard error), median redshift  $\tilde{z}$  and standard deviation  $\sigma$  of the distribution is given.

Field	$\bar{z}$	$\tilde{z}$	$\sigma$
Q2359	$2.81 \pm 0.03$	2.74	0.36
Q0301	$2.64 \pm 0.02$	2.59	0.31
Q2231	$2.68 \pm 0.03$	2.65	0.30
HE 0940	$2.79 \pm 0.02$	2.77	0.34
Q2348	$2.90 \pm 0.02$	2.92	0.36
Total	$2.79 \pm 0.01$	2.76	0.35



**Figure 12.** Redshift distribution of all identified sources in our five target fields, separated by initial candidate priority. Panel (a) shows the LBG\_PRI1, LBG\_PRI2 and LBG\_PRI3 classes, while the LBG\_DROPs are shown separately in panel (b) for clarity. Colours are as in Figs 4–7.

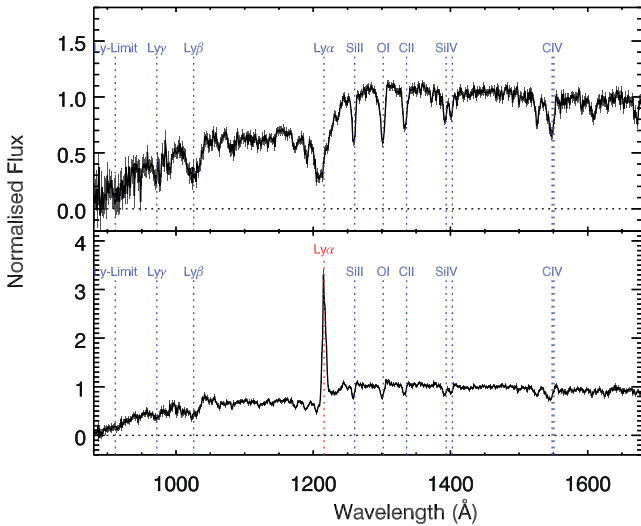


**Figure 13.** The distribution of redshift measurement errors,  $\Delta z$ , calculated using the LBGs which were observed twice in our survey and therefore have two independent redshift measurements. In panel (a) we show the full distribution, in panel (b) a close-up of the distribution at  $z < 0.025$ . Overplotted as a hashed histogram in panel (a) are the  $\Delta z$  values for three sources in our survey which had a redshift in the survey of Steidel et al. (2003).

disproportionately affects sources with an ID quality parameter  $Q = 0.5$ : of the four sources with large  $\Delta z$ , one was given a quality factor of 0.5 for both redshift measurements, while the other three have one measurement with  $Q = 0.5$  and another with  $Q = 0.6$ ; in the latter cases the  $Q = 0.6$  measurement is fairly robust while the  $Q = 0.5$  measurement is less reliable. Therefore, the LBGs which may suffer from large errors can be excluded by removing the  $Q = 0.5$  LBGs from the sample.

### 2.7.1 Composite spectra

We have calculated composite spectra using the  $z > 2$  VIMOS low-resolution galaxy data. The composite spectra were generated by



**Figure 14.** Composite spectra for galaxies showing  $\text{Ly}\alpha$  in absorption (top) and emission (bottom), with ISM absorption lines clearly detected.

averaging over the spectra after having corrected the spectra for the instrument response and having masked skylines. In addition, each individual spectrum is normalized by its median flux in the rest-frame wavelength range  $1250 \leq \lambda \leq 1450 \text{ \AA}$  before being combined to form the composite.

In Fig. 14 we show stacked spectra for the LBGs, separated into those showing  $\text{Ly}\alpha$  in emission (lower panel) and in absorption (upper panel). These stacked spectra show the average UV spectral energy distribution (SED) of a  $z \approx 3$  LBG with excellent SNR, and the quality of these spectra provide an indication of the robustness of our LBG identifications.

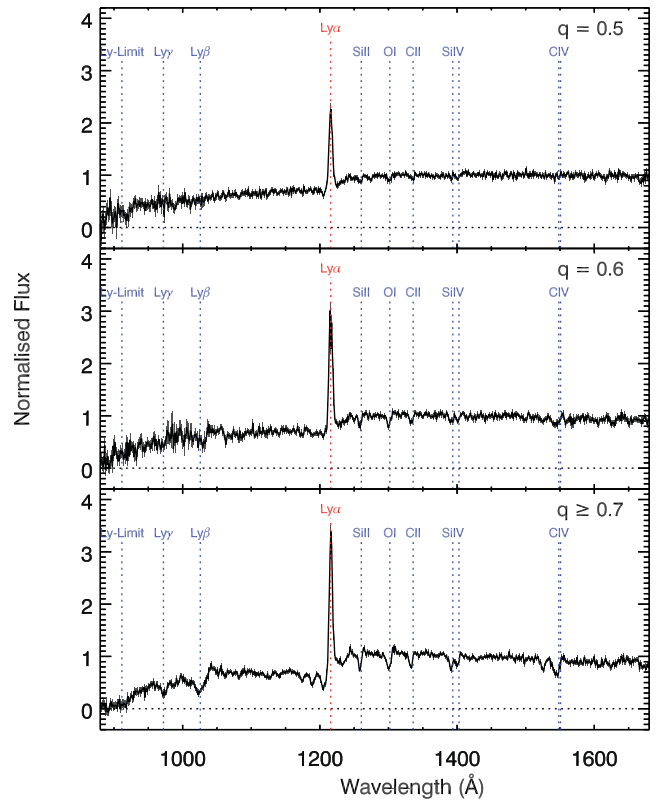
In Fig. 15, we show three separate composite spectra for sources classified as LBe's with quality IDs of  $Q = 0.5$ ,  $Q = 0.6$  and  $Q \geq 0.7$ . These spectra reflect the quality criteria set out in Section 2.5 well, with increasing quality spectra clearly showing increasingly high SNR in both  $\text{Ly}\alpha$  emission and ISM absorption features. In addition, the strength of the absorption features in the spectra appears to be systematically weaker with lower  $Q$ . This is likely the result of the lower SNR of the lower  $Q$  identifications.

Finally, we note that some potential flux is observed at wavelengths below the Lyman limit, however, even after stacking, the signal is subject to significant noise. Further analysis on the escape fraction may be possible using this data, but is beyond the scope of this paper.

### 2.7.2 Quasars and AGN

We have identified 33  $z > 1.5$  AGN and QSOs in our spectroscopic sample, which we present here as part of the VLRS catalogue. They have been identified by the presence of strong  $\text{Ly}\alpha$ ,  $\text{CIV}$  and  $\text{CIII}+\text{SiIII}$  emission lines as well as the generally weaker lines of  $\text{OVI}$ ,  $\text{NV}$  and  $\text{SiIV}$ . Their spectra are shown in Fig. 16. The above emission lines are indicated in each panel of Fig. 16.

Several other emission lines are detected in some of the QSO spectra but are not marked in the figure.  $\text{Ly}\beta$   $\lambda 1026$  is clearly seen in panel (e), where it may be asymmetrically broadened to longer wavelengths by the presence of relatively weak  $\text{OVI}$   $\lambda 1035$ . Panel (p) shows an emission line peaking at  $1029 \text{ \AA}$ , likely suggesting a blend between  $\text{Ly}\beta$  and  $\text{OVI}$ .



**Figure 15.** Composite spectra for galaxies classified as LBe's, separated by ID quality parameter  $Q$ . Those with  $Q = 0.5$  (upper panel) show comparatively weak  $\text{Ly}\alpha$  emission and marginal detections of ISM absorption lines. Moving to higher quality values, the strength of the  $\text{Ly}\alpha$  line increases and we also see high-SNR absorption features. Note that the spectra are not flux calibrated.

Emission arising from the combination of  $\text{OI}$   $\lambda 1302$  and  $\text{SiII}$   $1304$  is visible in a number of the spectra, for example in panels (a) and (h). Finally, many of the panels show clear emission at  $\approx 1400 \text{ \AA}$ , arising from a blend of the  $\text{SiIV}$   $\lambda 1396$  and  $\text{OVI}$   $\lambda 1402$  transitions.

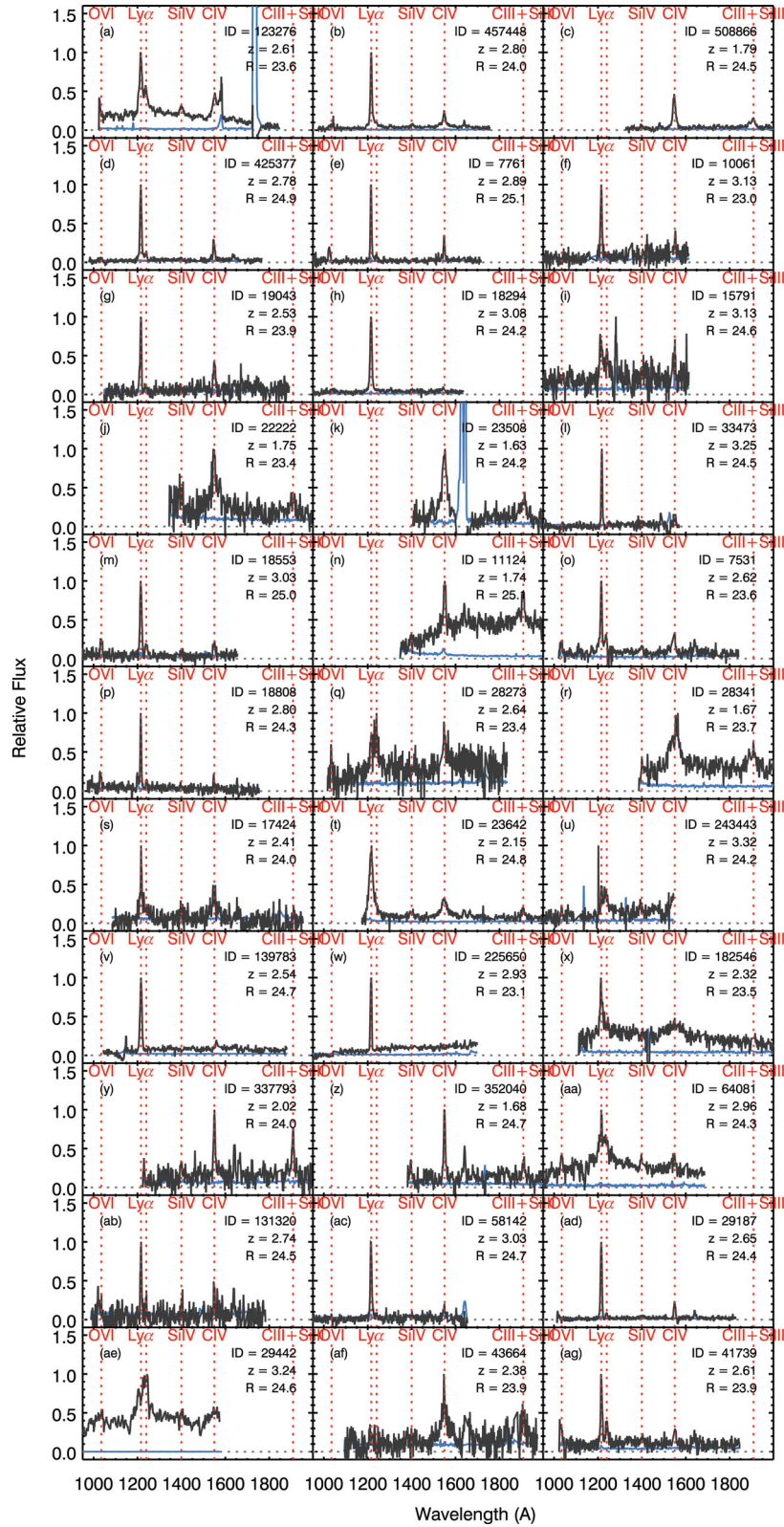
The spectra show a clear mix of both broad and narrow line AGN, the narrow line objects suggestive of the presence of obscured AGN activity. These are reminiscent of the AGN identified in similar  $z \sim 3$  star-forming galaxy surveys, for example Steidel et al. (2002) and Hainline et al. (2011, 2012).

The redshift distribution of the 33  $z > 1.5$  AGN is shown in the left-hand panel Fig. 17, whilst the  $R$ -band Vega magnitude distribution is shown in the right-hand panel.

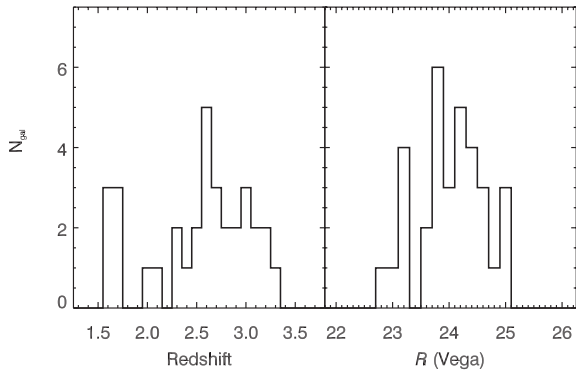
## 3 CLUSTERING OF $z \approx 3$ LBGs

We now analyse the clustering of the  $z \approx 3$  LBGs. As well as offering insights into the growth and evolution of structure in the Universe, we aim to measure the dynamics of the  $z \approx 3$  galaxy population, i.e. peculiar velocities and gravitational infall, to inform the analysis of the gas–galaxy relationship via LBG– $\text{Ly}\alpha$  forest cross-correlation (see Paper II).

We note that for the purposes of the clustering analysis we use the 1994  $R \leq 25$  VLRS sample (and place a limit of  $R \leq 25.2$  on the Keck sample with which it is compared and combined). Aside from this magnitude cut, all galaxies with  $Q \geq 0.5$  are included throughout this analysis. Taking the  $R = 25.2$  limit for the Keck



**Figure 16.** Rest-frame VIMOS spectra for the QSOs detected in the five fields presented. The y-axis scales differ from panel to panel, but the dotted line marks zero flux in each case. Dashed red lines indicate the wavelength of O VI, Ly $\alpha$ , Ne V and C IV emission. Gaps in the spectra indicate that an artefact has been masked out.



**Figure 17.** The redshift (left) and  $R_{\text{Vega}}$  magnitude (right) distributions of the 33 serendipitous spectroscopically identified quasar/AGN sample.

sample provides 815 galaxy redshifts within the Steidel et al. (2003) fields to combine with our VLRS sample.

In the analysis that follows, we measure the galaxy clustering as a function of galaxy–galaxy separation using the Landy–Szalay estimator:

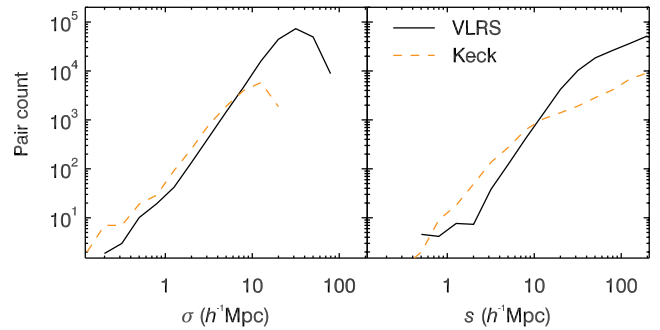
$$\xi(x) = \frac{\langle \text{DD}(x) \rangle - 2 \langle \text{DR}(x) \rangle + \langle \text{RR}(x) \rangle}{\langle \text{RR}(x) \rangle}, \quad (3)$$

where  $\xi(x)$  is the clustering as a function of separation  $x$ ,  $\text{DD}(x)$  is the number of galaxy–galaxy pairs at that separation,  $\text{DR}(x)$  is the number of galaxy–random pairs and  $\text{RR}(x)$  is the number of random–random pairs. This is estimated using a random catalogue that consists of  $20\times$  as many random points as data points and that covers an identical area. The redshift distribution of the random catalogue is set using a polynomial fit to the data.

We focus on fitting the autocorrelation function in the semiprojected,  $w_p(\sigma)$ , and full 2D,  $\xi(\sigma, \pi)$ , forms, where  $\sigma$  and  $\pi$  are the separation of two galaxies transverse and parallel to the line-of-sight. But we shall also study the LBG  $z$ -space correlation function,  $\xi(s)$ , where the signal can be significantly higher at large scales.

For  $\xi(\sigma, \pi)$  in particular, we also consider the combined sample of the VLRS data with the Keck LBG data of Steidel et al. (2003). The Keck data offer higher sampling rates than the VLRS, but across smaller field sizes ( $\approx 8$  arcmin). This is illustrated in Fig. 18 where the solid black line shows the VLRS pair counts (DD) as a function of separation in the transverse direction,  $\sigma$  (left-hand panel), and the 3D separation,  $s = \sqrt{\sigma^2 + \pi^2}$  (right-hand panel). In both panels the Keck pair counts are shown by the dashed orange line. Fig. 18 shows that the VLRS pair counts in both the transverse and 3D distance are significantly higher than in the previous Keck sample at  $\sigma, s \gtrsim 10 h^{-1} \text{ Mpc}$ .

On the validity of combining our own LBG sample with that of Steidel et al. (2003), we note that Steidel et al. (2003) used photometry with mean  $1\sigma$  depths of  $\langle \sigma(U_n) \rangle = 28.3$ ,  $\langle \sigma(G) \rangle = 28.6$  and  $\langle \sigma(R) \rangle = 28.0$ , whilst their imposed  $\mathcal{R}$ -band limit was  $\mathcal{R} = 25.5$ . Using the transformations of Steidel & Hamilton (1993), the Steidel et al. (2003)  $1\sigma$  limits correspond to  $U = 27.55$ ,  $B = 28.77$  and  $R = 27.86$  in the Vega system. Comparing this to the average depths across all the VLRS fields, we have mean  $3\sigma$  depths of  $U = 25.8$ ,  $B = 26.4$  and  $R = 25.9$ , which equate to  $1\sigma$  depths of  $U = 27.0$ ,  $B = 27.6$  and  $R = 27.1$ , approximately 1 mag fainter in each band than the Steidel et al. (2003) imaging data. However, given our limit of  $R = 25$  (and our imposed  $\mathcal{R} = 25.2$  for the Keck sample), the  $B - R$  (and  $G - \mathcal{R}$ ) constraints on the selections, and the inclusion of galaxies with no  $U$  detection, the VLRS and



**Figure 18.** Pair counts (DDs) as a function of pair separation in the transverse,  $\sigma$ , and 3D Hubble,  $s$ , directions. The Keck data (dashed orange curves) provide better sampling at small scales, whilst the VLRS data (solid black curves) provide sampling at larger scales that is not provided by previous data.

Keck samples will be relatively equivalent in terms of the galaxies included.

It is clear, however, that the VLRS sample, although giving a close approximation of the Keck sample selections, is not a perfect replica of the Keck selection. Given the difference in the filters and the moderate difference in depths this was unlikely to ever be the case. The redshift distributions are relatively well matched, but (partially due to the addition of the LBG\_PRI3 selection) the VLRS sample is skewed somewhat to marginally lower redshifts (as illustrated in Fig. 11). Additionally, the sky and space densities are close but not perfectly matched, as are the  $R$ -band magnitude distributions as shown by Paper I. As a result, the UV luminosity functions will be similarly close but not perfectly matched. In combining the two samples we therefore note these differences and use the results of combining the samples with caution. However, it is beneficial to do so in order to help constrain the redshift-space distortion effects, which are an important element of further work incorporating the Ly $\alpha$  forest to constrain the distribution of gas around these star-forming galaxies. Furthermore, it is difficult to perform these tests with either sample alone given the Keck sample's small area coverage and the VLRS sample's comparatively lower sampling rate. Therefore, although the combination is not ideal, it offers an indication of the impact of redshift-space distortions on the correlation functions that may be utilized in subsequent work.

### 3.1 Semiprojected correlation function, $w_p(\sigma)$

We first estimate the LBG clustering using the semiprojected correlation function,  $w_p(\sigma)$ . This gives the clustering at fixed transverse separation,  $\sigma$ , integrated over line-of-sight distance,  $\pi$ , approximately independent of the effect of peculiar velocities and is given by

$$w_p(\sigma) = 2 \int_0^\infty \xi(\sigma, \pi) d\pi, \quad (4)$$

where  $\xi(\sigma, \pi)$  is the 2D autocorrelation function. We integrate  $\xi(\sigma, \pi)$  over the range  $0 < \pi < \pi_{\text{max}}(\sigma)$ , where  $\pi_{\text{max}}(\sigma)$  is given by the maximum of  $1000(1+z)/H(z)$  and  $15\sigma$  at a given sky separation  $\sigma$  (consistent with Adelberger et al. 2003; da Ângela et al. 2005b).

In the calculation of  $w_p(\sigma)$ , we make a correction for the effect of ‘slit collisions’, following Paper I. Any object observed with VIMOS takes up an area on the detector of at least  $40 \times 640$  pixels (Section 2.4.1), corresponding to  $8 \times 130 \text{ arcsec}^2$  on-sky. Other candidates lying within this area can therefore not be observed (unless the area is revisited), and as a result, pairs of LBGs at small



separations are systematically missed by our survey. This effect will reduce the measured LBG autocorrelation at small separations. Paper I quantified this effect by comparing the angular autocorrelation function of photometrically selected LBG candidates and spectroscopically observed candidates. Using their result, we correct for this effect in our LBG survey by weighting DD pairs at  $\theta < 8$  arcmin according to the weighting factor given by

$$W(\theta) = \frac{1}{1 - (0.0738 \theta^{-1.052})}, \quad (5)$$

where  $\theta$  is the angular separation in arcmin.

In addition to the slit collision correction, a further correction – the integral constraint – is required to compensate for the effect of the limited field sizes. For this we follow the commonly used approach of using the random–random pair distributions, which have been constructed to match the survey geometry, to determine the magnitude of the integral constraint. This method has been well described by a number of authors (e.g. Groth & Peebles 1977; Peebles 1980; Roche et al. 1993; Baugh et al. 1996; Phleps et al. 2006), with Phleps et al. (2006) in particular providing a detailed discussion in relation to the projected correlation function, and we provide a brief description of the calculation here.

The measured correlation function is given by the true correlation function minus the integral constraint,  $\mathcal{I}$ :

$$w_m(\sigma) = w_t(\sigma) - \mathcal{I}. \quad (6)$$

Assuming a power law form for the real-space clustering, the true projected clustering is fit by

$$w_p(\sigma) = C r_0^\gamma \sigma^{1-\gamma}, \quad (7)$$

where  $r_0$  is the real-space clustering length and  $\gamma$  is the slope of the real-space clustering function,  $\xi(r)$ , which is characterized by a power law of the form

$$\xi(r) = \left(\frac{r_0}{r}\right)^\gamma. \quad (8)$$

The factor  $C$  is given by

$$C = \left( \frac{\Gamma(\frac{1}{2}) \Gamma(\frac{\gamma-1}{2})}{\Gamma(\frac{\gamma}{2})} \right), \quad (9)$$

where  $\Gamma$  is the Gamma function. Given this framework, the integral constraint can be estimated from the mean of the random–random pair counts,  $\langle RR \rangle$ , and the slope of the correlation function, such that

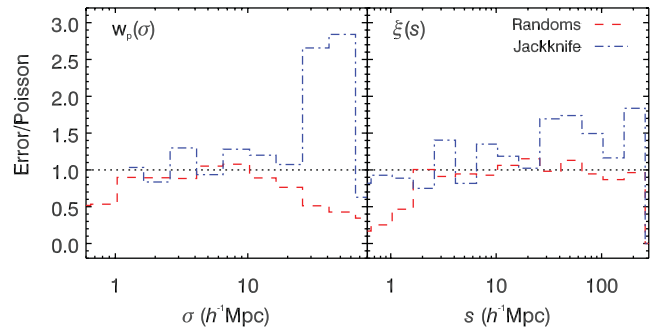
$$\frac{\mathcal{I}}{C r_0^\gamma} = \frac{\Sigma \langle RR(\sigma) \rangle \sigma^{1-\gamma}}{\Sigma \langle RR(\sigma) \rangle}. \quad (10)$$

Quantifying errors on the autocorrelation function has been performed using Poissonian, jack-knife and random realization error estimates. The Poisson errors are given by

$$\Delta \xi = \frac{(1 + \xi)}{\sqrt{\langle DD \rangle / 2}}. \quad (11)$$

The jack-knife errors were computed by splitting the data into individual fields, with the large fields (i.e. HE 0940 and Q2348) being split into two fields each. We therefore have 11 different jack-knife realizations with a single field (or half-field) being excluded in each realization.

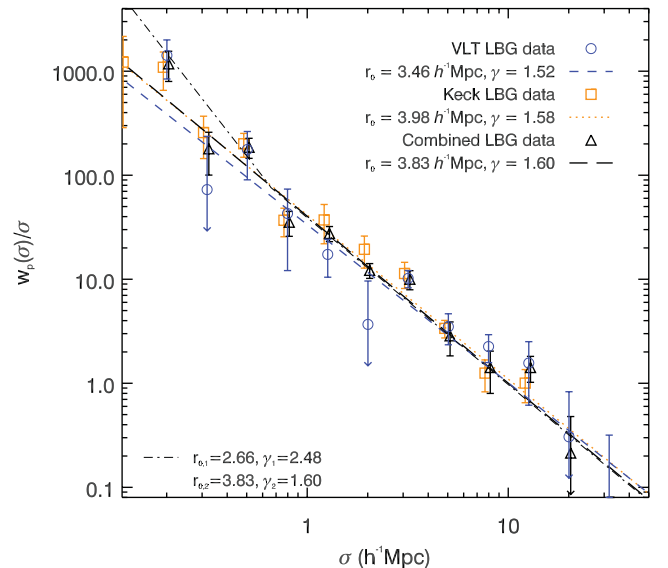
The random realization error estimates incorporate 100 random catalogues with the same number of objects as the real data. We then calculate the correlation function using these random realizations to calculate the  $\langle DD \rangle$  pairs and take the standard deviation of the results as the uncertainty on the measurement.



**Figure 19.** Ratio of the random mock (red dashed histogram) and the jack-knife (blue dot-dashed histogram) error estimates to the Poisson errors as a function of separation for the projected correlation function (left-hand panel) and the redshift-space autocorrelation function (right-hand panel). The results are shown for the VLRS sample only.

In Fig. 19, we compare the above error estimates for  $w_p(\sigma)$  and the redshift-space clustering function,  $\xi(s)$  (see Section 3.2), showing the ratio of the jack-knife and random realization methods to the Poisson result. The estimates are consistent with each other over scales from  $\approx 1$ – $25 h^{-1}$  Mpc. In what follows, we therefore use the Poisson estimates at separations of  $< 25 h^{-1}$  Mpc and jack-knife estimates at separations  $> 25 h^{-1}$  Mpc.

The projected autocorrelation function for the VLRS sample (black circles), the Keck sample (orange squares; Steidel et al. 2003) and the two combined (cyan triangles) is shown in Fig. 20. The plotted data include the integral constraint correction, which



**Figure 20.** Projected autocorrelation function,  $w_p(\sigma)/\sigma$ , for the VLT LBG sample (black circles). The result for the  $2.5 < z < 3.5$  Keck LBG sample is shown by the orange squares and error bars for both results are estimated using Poisson errors. The best-fitting power laws for both the VLT (solid black line) and Keck (dashed orange line) are shown, with clustering amplitudes of  $r_0 = 3.46 \pm 0.41$  and  $3.98 \pm 0.32 h^{-1}$  Mpc, with slopes of  $\gamma = 1.52 \pm 0.13$  and  $1.58 \pm 0.13$ , respectively. The cyan triangles show the combined result and the dotted cyan line the best fit to this of  $r_0 = 3.83 \pm 0.24 h^{-1}$  Mpc with a slope of  $\gamma = 1.60 \pm 0.09$ . The black dash-dot line shows a double power law, motivated by Paper I, with  $r_0 = 2.61 h^{-1}$  Mpc and  $\gamma = 2.48$  below the break and  $r_0 = 3.75 h^{-1}$  Mpc and  $\gamma = 1.61$  above it.



we estimated as  $\mathcal{I}_{\text{wp}} = 5.33$  and  $7.18$  for the VLRS and Keck data, respectively.

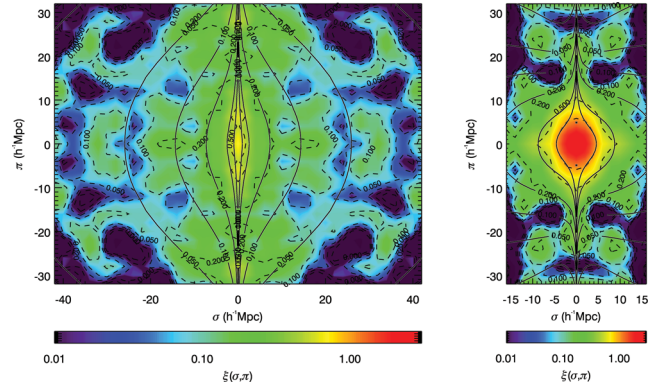
Based on this we estimate a clustering length of the entire VLRS sample of  $r_0 = 3.46 \pm 0.41 h^{-1} \text{ Mpc}$  (comoving) with a slope of  $\gamma = 1.52 \pm 0.13$ . The Keck result on its own gives a result of  $r_0 = 3.98 \pm 0.32 h^{-1} \text{ Mpc}$  with  $\gamma = 1.58 \pm 0.13$ , whilst the combined VLRS+Keck data gives  $r_0 = 3.83 \pm 0.24 h^{-1} \text{ Mpc}$  with a slope of  $\gamma = 1.60 \pm 0.09$ . These  $r_0$  results are comparable to the clustering of star-forming galaxies at lower redshifts (e.g. Blake et al. 2009; Bielby et al. 2010).

Comparing to other measurements of the  $z \approx 3$  LBG clustering length, Giavalisco & Dickinson (2001) measured  $r_0 = 5.0 \pm 0.7 h^{-1} \text{ Mpc}$  for  $R_{\text{AB}} < 25$  LBGs, but for a relatively small number of galaxies ( $\approx 400$ ). Building on that sample, Adelberger et al. (2003) measured  $r_0 = 3.96 \pm 0.29 h^{-1} \text{ Mpc}$  with a slope of  $\gamma = 1.55 \pm 0.15$  at  $R_{\text{AB}} < 25.5$ . We note that with the same sample, but a different method, Adelberger et al. (2005a) found a higher clustering strength of  $r_0 = 4.5 \pm 0.6 h^{-1} \text{ Mpc}$ . Subsequently, Cooke et al. (2006) measured the clustering of  $z \approx 3$  LBGs in fields around damped Ly $\alpha$  absorbers and found a lower clustering strength of  $r_0 = 2.65 \pm 0.48 h^{-1} \text{ Mpc}$  with a slope of  $\gamma = 1.55 \pm 0.40$  at  $R_{\text{AB}} < 25$ , whilst Trainor & Steidel (2012) performed a similar measurement but around  $z \sim 2.7$  QSOs (and with a galaxy sample incorporating a mixture of LBG and BX selections) and found a clustering length of  $r_0 = 6.0 \pm 0.5 h^{-1} \text{ Mpc}$ . Overall, our result appears consistent with other measurements, although marginally lower than the Giavalisco & Dickinson (2001) and Adelberger et al. (2003, 2005a) results, which are all based on the same – or a subset of the same – sample. As observed by some of the above authors, the LBG clustering lengths are generally somewhat smaller than those measured for the slightly lower redshift BM and BX selections.

The above estimates are based on spectroscopically confirmed samples and a number of clustering measurements exist based on photometric samples. For example Foucaud et al. (2003) measured  $r_0 = 5.9 \pm 0.5 h^{-1} \text{ Mpc}$  from the angular correlation function of  $R_{\text{AB}} < 24.5$  LBGs from the Canada–France Deep Fields Survey (McCracken et al. 2001), a higher  $r_0$  than the spectroscopic samples, but also a significantly brighter magnitude cut. Additionally, Adelberger et al. (2005a) measured  $w(\theta)$  for photometrically selected LBGs and found  $r_0 = 4.0 \pm 0.6 h^{-1} \text{ Mpc}$  for  $R \leq 25.5$  LBGs, consistent with our results. Hildebrandt et al. (2007) measured  $r_0 = 4.8 \pm 0.3 h^{-1} \text{ Mpc}$  for  $22.5 < R_{\text{Vega}} < 25.5$  galaxies in the Garching–Bonn Deep Survey (GaBoDS) data. Subsequently, Hildebrandt et al. (2009) measured  $r_0 = 4.25 \pm 0.13 h^{-1} \text{ Mpc}$  for Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) LBGs at  $r_{\text{AB}} < 25$  and using photo- $z$  from HYPERZ (Bolzonella, Miralles & Pelló 2000). In general, the clustering measured from photometric samples appears to give somewhat larger clustering lengths than those obtained for the spectroscopic samples. As with our own sample, however, these selections are not perfect replicas of the original  $U_n \text{GR}$ -based selection and this may be part of the cause of this, perhaps resulting in subtle differences in the redshift or luminosity ranges.

### 3.2 2D autocorrelation function, $\xi(\sigma, \pi)$

As discussed above, integrating along the redshift/line-of-sight direction leaves the semiprojected correlation function,  $w_p(\sigma)$ , independent of the effects of galaxy peculiar motions. We now attempt to fit the full 2D correlation function,  $\xi(\sigma, \pi)$ , to retrieve



**Figure 21.** The two-dimensional autocorrelation function,  $\xi(\sigma, \pi)$  results for the VLRS (left) and Keck (right) data samples individually. The shaded contour map and dashed contour lines show the measured  $\xi(\sigma, \pi)$  in each case.

the kinematics of the galaxy population and to make new estimates of  $r_0$ .

As before, we use the Landy–Szalay estimator to calculate the correlation function but now as a function of both transverse separation,  $\sigma$ , and line-of-sight separation,  $\pi$ . We use the same random catalogues matching the survey fields as used for the calculation of the projected correlation function. We again calculate the integral constraint for the data sets using the random catalogues via

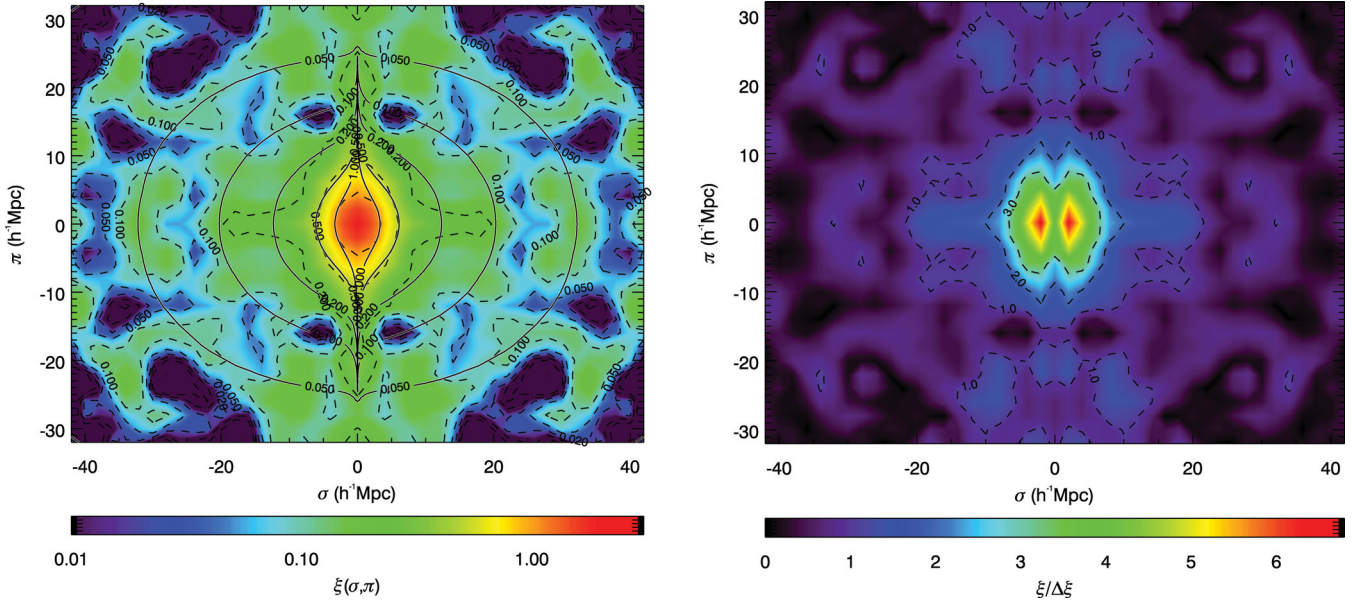
$$\frac{\mathcal{I}}{r_0^\gamma} = \frac{\Sigma \langle \text{RR}(s) \rangle s^\gamma}{\Sigma \langle \text{RR}(s) \rangle}, \quad (12)$$

where  $s = \sqrt{\sigma^2 + \pi^2}$ . This gives values of  $\mathcal{I}_\xi = 0.024$  and  $0.064$  for the VLRS and Keck data samples, respectively.

Fig. 21 shows the result for the VLRS data (left-hand panel), which provides a greater handle on the large scale ( $s \gtrsim 10 h^{-1} \text{ Mpc}$ ) clustering, the Keck data (right-hand panel), which provides greater sampling on small scales. Fig. 22 shows the VLRS and Keck results combined. In each case,  $\xi(\sigma, \pi)$  was calculated in linear  $2 h^{-1} \text{ Mpc}$  bins and subsequently smoothed with a FWHM of  $2 h^{-1} \text{ Mpc}$ .

For both the VLRS and Keck samples, we see the ‘finger-of-God’ effect at small  $\sigma$  scales in which the clustering power is extended in the  $\pi$  direction. This effect is a combination of galaxy peculiar velocities and measurement errors on the galaxy redshifts. In addition, in the VLRS a flattening of the clustering measurement at large scales is evident, which is caused by dynamical infall of galaxies.

We now fit models of the clustering to these results, initially assuming a single power law for  $\xi(r)$  and allowing  $r_0$  and the kinematical parameters to vary. We take the  $r_0$  and  $\gamma$  estimates from the  $w_p(\sigma)$  fit as the starting point in fitting the 2D clustering. The kinematics are characterized by two parameters: the velocity dispersion in the line of sight direction  $\sqrt{\langle w_z^2 \rangle}$  and the infall parameter,  $\beta$ . The model we use incorporating the galaxy kinematics is described in full by Hawkins et al. (2003) and Paper I. The model accounts for two key effects on the clustering statistics caused by galaxy motions. The first is the finger-of-God effect, which is constrained by the velocity dispersion and the second is the Kaiser effect (i.e. the coherent motion of galaxies on large scales), which is characterized by  $\beta$ .

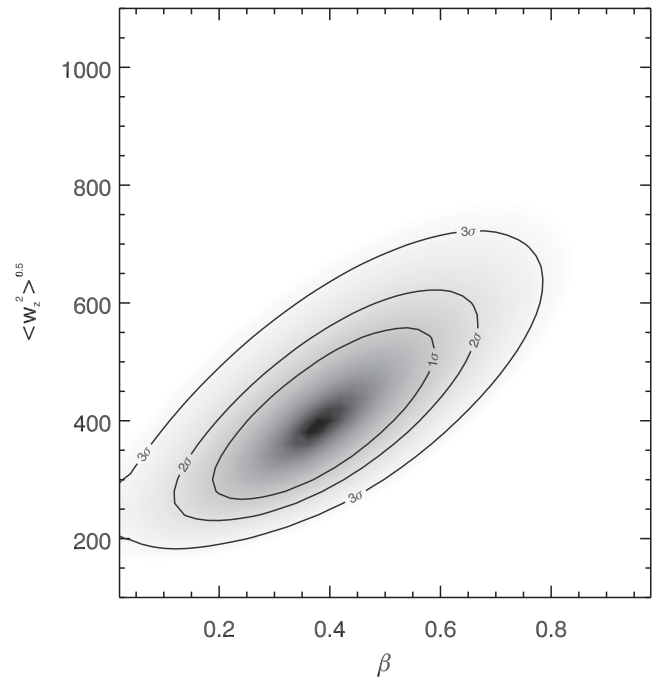


**Figure 22.** The two-dimensional autocorrelation function,  $\xi(\sigma, \pi)$ , results for the VLRS sample (left) and the SNR on the result (right). The shaded contour map and dashed contour lines show the measured  $\xi(\sigma, \pi)$ , whilst the solid contour lines give the best-fitting model.

For the VLRS and the combined samples we fit over the range  $1.0 \leq s \leq 25 h^{-1}$  Mpc, whilst for the Keck data by itself we limit the fit to the scales  $1.0 \leq s \leq 15 h^{-1}$  Mpc (note that the largest single field available in the Keck data is  $\approx 15 h^{-1}$  Mpc).

For the two samples individually, we find that it is difficult to place reasonable constraints on both the velocity dispersion and the infall together. With the VLRS data (over the range  $1 \leq (\sigma, \pi) \leq 25 h^{-1}$  Mpc), we find  $\beta(z=3) = 0.3^{+1.7}_{-0.3}$  and  $\sqrt{\langle w_z^2 \rangle} = 1700^{+2000}_{-900} \text{ km s}^{-1}$ , the low SNR on small scales limiting the fit accuracy. We experimented with adding a uniform error distribution out to  $\pm 12\,000 \text{ km s}^{-1}$  to the Gaussian velocity dispersion (cf. Fig. 13) but this made little difference in the  $\sigma, \pi$  range fitted. Fitting the Keck data gives best-fitting values of  $\beta(z=3) = 0.85^{+0.30}_{-0.35}$  and  $\sqrt{\langle w_z^2 \rangle} = 700 \pm 220 \text{ km s}^{-1}$ . We note that da Ângela et al. (2005b) performed a similar fit to the Keck data for  $\beta(z=3)$ , but kept a constant velocity dispersion of  $\sqrt{\langle w_z^2 \rangle} = 400 \text{ km s}^{-1}$ , finding a value for the infall parameter of  $\beta(z=3) = 0.15^{+0.20}_{-0.15}$ . By also setting the velocity dispersion to a value of  $400 \text{ km s}^{-1}$ , we find that we retrieve a comparable result to da Ângela et al. (2005b), highlighting the degeneracy between the velocity dispersion and the infall parameter.

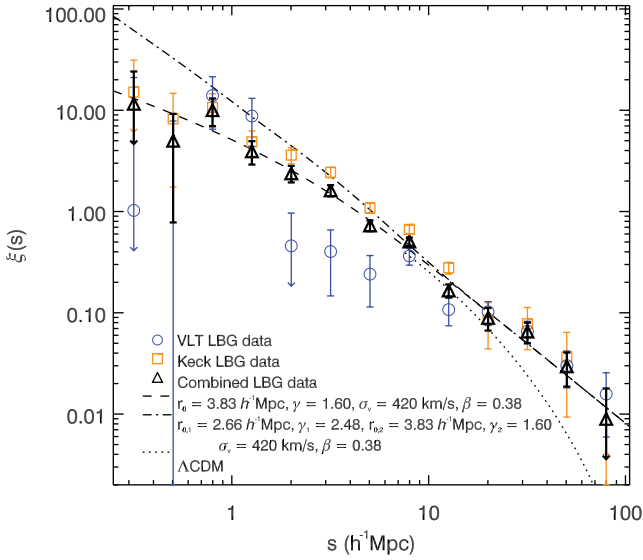
Ultimately, fitting the VLRS  $\xi(\sigma, \pi)$  is hindered by a lack of SNR on small scales, whilst the fit to the Keck data is hindered by the small size of the fields. We thus combine the two data sets and fit the full LBG sample in the same manner as with the individual samples. The fit is performed in the range  $1 < s < 25 h^{-1}$  Mpc and we allow the velocity dispersion and the infall parameter to vary. The resulting fit gives a velocity dispersion of  $\sqrt{\langle w_z^2 \rangle} = 420^{+140}_{-160} \text{ km s}^{-1}$  and an infall parameter of  $\beta(z=3) = 0.38 \pm 0.19$ . We show the  $\chi^2$  contours for the fit in the  $\beta(z=3) - \sqrt{\langle w_z^2 \rangle}$  plane in Fig. 23 (the contours represent the  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  confidence limits). From this figure, the degeneracy can be seen between  $\sqrt{\langle w_z^2 \rangle}$  and  $\beta$ , where increasing  $\beta$  similarly increases the best-fitting velocity dispersion. The best-fitting results are plotted over the contour maps of the  $\xi(\sigma, \pi)$  measurements in Fig. 22 (dashed contours). As with the data, we see the finger-of-god and large-scale flattening effects in the fitted models.



**Figure 23.** Fitting contours for peculiar velocity and bulk inflow based on the combined VLRS+Keck  $\xi(\sigma, \pi)$ . The best-fitting result is given by  $\beta(z=3) = 0.38 \pm 0.19$  and  $\sqrt{\langle w_z^2 \rangle} = 420^{+140}_{-160} \text{ km s}^{-1}$ .

### 3.3 Redshift-space correlation function, $\xi(s)$

In order to check the consistency of our measurements, we now compare the model fit obtained from  $w_p(\sigma)$  and  $\xi(\sigma, \pi)$  to the measured redshift-space autocorrelation function  $\xi(s)$ . Again we use the Landy–Szalay estimator and quote errors based on Poisson estimates. The  $\xi(s)$  results for the VLRS, Keck and combined LBG samples are shown in Fig. 24. We also plot the single power-law estimate of the intrinsic clustering from our fits to the VLT+Keck



**Figure 24.** Redshift-space correlation function,  $\xi(s)$ , for the VLRS (open blue circles), Keck (orange squares) and combined (black triangles) samples. The short and long dashed curves show single power-law model fits to the combined sample based on fits to  $w_p(\sigma)$  and  $\xi(\sigma, \pi)$ . The dash-dot curve gives a double power-law model fit, whilst the dotted curve at large scales shows the predicted  $\Lambda$ CDM clustering.

$\xi(\sigma, \pi)$  (dotted line) and the result of this power law after applying the best-fitting values for  $\beta$  and  $\sqrt{\langle w_z^2 \rangle}$  (dashed line). The final fit incorporating the galaxy dynamics is marginally low compared to the data points, but is easily consistent within the error bars.

Our measurements of  $\beta$  and  $\sqrt{\langle w_z^2 \rangle}$  are consistent with the previous measurements using the first VLT data set (Paper I,  $\beta = 0.48$  and  $\sqrt{\langle w_z^2 \rangle} = 700 \text{ km s}^{-1}$ ). As discussed in Paper I, the median measurement error on the galaxy lines on the VLT VIMOS spectra is  $\approx 350 \text{ km s}^{-1}$ . In addition, an uncertainty of  $\approx 100 \text{ km s}^{-1}$  is introduced by the transformation from outflow redshifts to intrinsic galaxy redshifts (Steidel et al. 2010). The final contribution to the velocity dispersion is from the intrinsic peculiar velocities of the galaxies. Using the GIMIC simulations (Crain et al. 2009) we have analysed the mean velocity dispersion of LBG-like galaxies and find a value of  $\approx 100 \text{ km s}^{-1}$ . Combining these three elements in quadrature, we would expect a pairwise velocity dispersion of  $\sqrt{\langle w_z^2 \rangle} \approx \sqrt{2} \times \sqrt{(350 \text{ km s}^{-1})^2 + (100 \text{ km s}^{-1})^2 + (100 \text{ km s}^{-1})^2} \approx 500 \text{ km s}^{-1}$ . This is within the  $1\sigma$  error contours given in Fig. 23. This value is also reasonably consistent with the VLRS  $\xi(s)$  estimate (see Fig. 24).

We also show in Fig. 24 the matter correlation,  $\xi(r)$ , scaled to the LBG clustering strength. This was calculated using the CAMB software and using a flat  $\Lambda$ CDM cosmology with  $\Omega_m = 0.27$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and  $\Omega_b h^2 = 0.022$ . There are currently some claims that non-Gaussianity is detected at  $z \approx 1$  in NRAO VLA Sky Survey radio source (Xia et al. 2010) and LRG data sets (Sawangwit et al. 2011; Thomas, Abdalla & Lahav 2011; Nikoloudakis, Shanks & Sawangwit 2012). The evidence generally comes from detecting large-scale excess power via flatter slopes for angular correlation functions. Since non-Gaussianity is easier to detect at high redshift this motivates checking the LBG  $\xi(s)$  for an excess. We have already noted that the slope from  $w_p$  and  $\xi(\sigma, \pi)$  at  $\gamma = 1.55$  is much flatter than the canonical  $\gamma = 1.8$ . This slope is also flatter than the  $z \approx 1$  LRG large-scale  $w(\theta)$  slope of Nikoloudakis et al. (2012). We

see that the VLRS does give reasonably accurate measurements for  $25 < s < 100 h^{-1} \text{ Mpc}$  and that the observed LBG  $\xi(s)$  shows a  $\approx 2.5\sigma$  excess over the  $\Lambda$ CDM model in this range. Even when the marginally smaller integral constraint for the  $\Lambda$ CDM model is assumed the discrepancy remains at  $\approx 2\sigma$ . We conclude that there is some evidence for an excess over the standard  $\Lambda$ CDM model but independent LBG data are needed to confirm this on the basis of the redshift-space correlation function. The statistical error on the LBG  $w(\theta)$  from Paper I is smaller but the flat power law here is only seen to  $\theta = 10 \text{ arcmin}$  or  $r = 13\text{--}14 h^{-1} \text{ Mpc}$  and this is not enough to decide the issue.

### 3.4 Double power-law correlation function models

We next look to see if a more complicated model than a power law for  $\xi(r)$  is required. This is motivated first because Paper I noted that there was an increase in the slope at  $\approx 1 h^{-1} \text{ Mpc}$  in the LBG angular autocorrelation function,  $w(\theta)$ , suggestive of the split between the one-halo and two-halo terms in the halo model of clustering. Although this result is uncertain due to quite significant low-redshift contamination corrections, such features have been seen in lower redshift galaxy samples, particularly for LRGs at  $z \approx 0.5$  (e.g. Ross et al. 2007; Sawangwit et al. 2011). Given the improved power of the VLRS, it is interesting to see if there is any evidence of a change in the slope at small scales in  $\xi(s)$  and  $w_p(\sigma)$  in our  $z \approx 3$  LBG sample.

We therefore show in Fig. 20 a double power-law model for  $w_p$  with the same power-law slopes as fitted by Paper I to the LBG  $w(\theta)$ . We have reduced the amplitude by  $\approx 20$  percent to match approximately the large-scale amplitude fitted to the VLRS and Keck combined data. This is within the systematic uncertainties of the  $w(\theta)$  measurement. Although certainly not required by the  $w_p$  data this double power law cannot be rejected by the combined  $w_p$  data, giving a reduced  $\chi^2$  of 1.77 (marginally smaller than the reduced  $\chi^2$  obtained for a single power law of 1.84).

In Fig. 24 we now compare to  $\xi(s)$  the same double power-law  $w(\theta)$  model with the  $\approx 20$  per cent reduced amplitude. Again with a velocity dispersion of  $420 \text{ km s}^{-1}$  and  $\beta = 0.38$  we see that the model cannot be rejected by the data. We note that if we use a  $\langle \text{DD} \rangle / \langle \text{DR} \rangle$  estimator the VLRS  $\xi(s)$  result shows increased power at large scales and the flatter slope of the double power-law model here provides a better fit.

We note that other authors have also reported a turn-up in the clustering at small scales in high-redshift galaxy samples. For instance Ouchi et al. (2005) report that  $z = 4$  LBG  $w(\theta)$  shows a steepening below  $\approx 0.2 h^{-1} \text{ Mpc}$  or  $\approx 10 \text{ arcsec}$  at  $z = 4$ . If both results are unaffected by contamination then it could argue for an evolutionary growth in this break scale between  $z = 4$  and 3.

Certainly there is plenty to motivate expanding surveys to make more accurate measurements of both the angular and redshift survey correlation functions at these redshifts. Below the break scale is of extreme interest for single halo galaxy formation models and at large scales the interest is in looking for a flattening of the correlation function slope due to the presence of primordial non-Gaussianity.

### 3.5 Estimating $\Omega_m$ and the growth rate

#### 3.5.1 The mass density of the Universe

We now look at the cosmological results afforded by the  $z \approx 3$  LBG clustering and dynamics. As discussed by Hoyle et al. (2002) and da Ângela et al. (2005a), it is, in principle, possible to constrain the

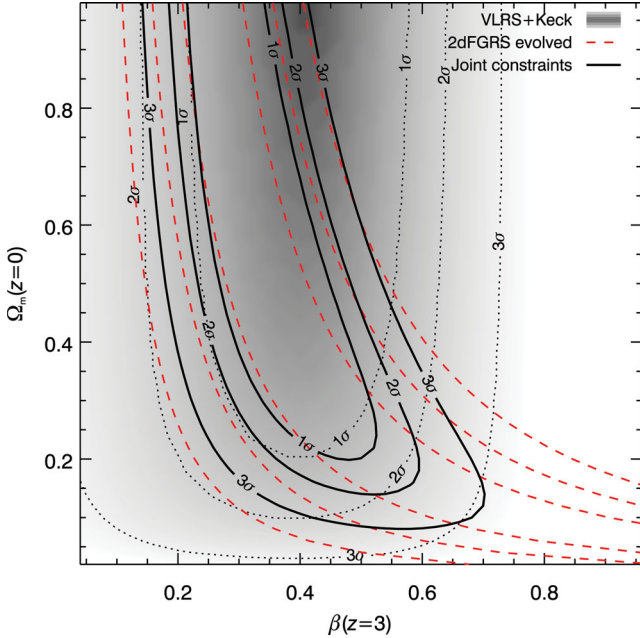


matter density  $\Omega_m(z=0)$  from the measurement of  $\xi(\sigma, \pi)$ . Effectively, the elongation of  $\xi(\sigma, \pi)$  along the line of sight increases with increasing values of  $\Omega_m(z=0)$ . However, increased values of  $\beta$  lead to a flattening of  $\xi(\sigma, \pi)$  along the line of sight. These effects combined lead to a degeneracy in determining  $\Omega_m$  from the galaxy clustering alone.

In previous sections, we have studied the galaxy dynamics assuming a cosmology with  $\Omega_m(z=0) = 0.3$ . We now fit the  $\xi(\sigma, \pi)$  result with this assumed cosmology, but now with a constant peculiar velocity of  $\sqrt{\langle w_z^2 \rangle} = 420 \text{ km s}^{-1}$  and fitting for  $\Omega_m(z=0)$  and  $\beta$ . The result is shown by the  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  contours (solid) in Fig. 25. Based on just the  $z=3$  galaxy clustering, we find results for the mass density of  $\Omega_m(z=0) = 0.08^{+0.22}_{-0.08}$  and on the infall parameter of  $\beta(z=3) = 0.38^{+0.15}_{-0.13}$ .

Breaking the degeneracy of this result can be achieved by incorporating lower redshift results as shown by da Ângela et al. (2005a, 2005b) and Paper I. As in these previous works, we use the 2dF Galaxy Redshift Survey (2dFGRS) measurements of Hawkins et al. (2003) to do this ( $r_0 = 5.0 h^{-1} \text{ Mpc}$ ,  $\gamma = 1.8$  and  $\beta(z=0.11) = 0.49 \pm 0.09$ ). The Hawkins et al. (2003) result can then be evolved to the redshift of our study based on the relationship between the growth parameter,  $f(z)$ , and the bulk motion and the clustering bias,  $b$ , of a galaxy population:

$$\beta = \frac{f(z)}{b} \approx \frac{\Omega_m(z)^{0.55}}{b}. \quad (13)$$



**Figure 25.** Fitting contours for the mass density and bulk inflow based on the combined VLRs+Keck  $\xi(\sigma, \pi)$ . The shaded region gives the result from the VLT+Keck data sample (assuming  $\sqrt{\langle w_z^2 \rangle} = 420 \text{ km s}^{-1}$ ,  $r_0 = 3.83 h^{-1} \text{ Mpc}$  and  $\gamma = 1.60$ ) with the dotted contour lines giving the  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  uncertainties. The results from this fit are  $\beta(z=3) = 0.38^{+0.15}_{-0.13}$  and  $\Omega_m(z=0) = 0.08^{+0.22}_{-0.08}$ . The dashed red lines show the  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  constraints given by evolving the 2dFGRS measurements as described in the text. The solid black contours give the combination of the two and give a result of  $\beta(z=3) = 0.38^{+0.16}_{-0.09}$  and  $\Omega_m(z=0) = 0.30^{+0.32}_{-0.18}$ .

The bias can be calculated directly from the clustering measurements by using the volume averaged clustering:

$$b = \sqrt{\frac{\xi_g(s)}{\xi_{DM}(s)}} = \sqrt{\frac{\bar{\xi}_g(8)}{\bar{\xi}_{DM}(8)}}, \quad (14)$$

where  $\bar{\xi}_g(8)$  is the volume averaged correlation function at  $s < 8 h^{-1} \text{ Mpc}$  for the galaxy population and  $\bar{\xi}_{DM}(8)$  is the same, but for the underlying dark matter distribution. The volume averaged clustering is calculated from the clustering using

$$\bar{\xi}(x) = \frac{3}{x^3} \int_0^x r^2 \xi(r) dr. \quad (15)$$

In addition, a measure of the dark matter clustering is required in order to estimate the bias of the galaxy population and we calculate this using the CAMB software incorporating the HALOFIT model of non-linearities (Smith et al. 2003). Using the previously determined best-fitting parameters of  $r_0 = 3.83 h^{-1} \text{ Mpc}$  and  $\gamma = 1.60$ , we evaluate the galaxy bias based on a single power law, finding a bias for the LBGs of  $b = 2.59 \pm 0.13$ .

We then determine the  $z=0.11$  underlying dark matter clustering amplitude from these parameter constraints and evolve this to  $z=3$  for test cosmology range of  $\Omega_m(z=0) = 0-1$ . The constraints on  $\beta$  using this method over a range of assumed  $\Omega_m$  values are given by the red dashed contours in Fig. 25. By combining these with the original constraints from  $\xi(\sigma, \pi)$ , we find a result of  $\beta(z=3) = 0.38^{+0.16}_{-0.09}$  and  $\Omega_m(z=0) = 0.30^{+0.32}_{-0.18}$ .

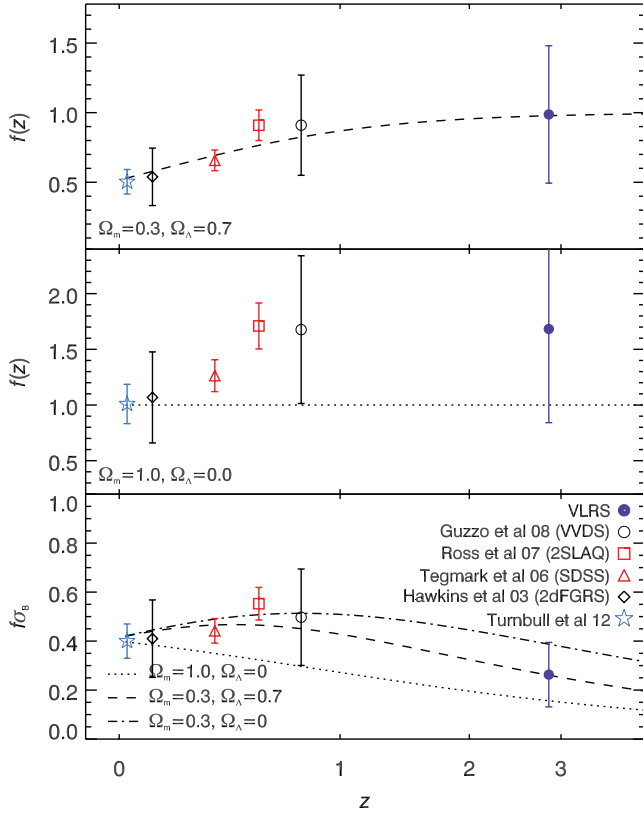
Across these analyses, we have consistently found a value for the infall parameter of  $\beta(z=3) \approx 0.36-0.40$ .  $\Omega_m$  is somewhat less well constrained, but remains consistent with  $\Lambda \text{CDM}$ . The measurements of  $\beta(z=3)$  presented here are consistent with our previous measurement from Paper I of  $\beta(z=3) = 0.48 \pm 0.17$ , whilst being somewhat higher than the result found by da Ângela et al. (2005b) of  $\beta(z=3) = 0.15^{+0.20}_{-0.15}$ . We note that the latter assumes a fixed velocity dispersion of  $\sqrt{\langle w_z^2 \rangle} = 400 \text{ km s}^{-1}$  and is limited to the small FoV of the Keck survey. As such, their lower estimate of  $\beta$  may well be a systematic of too small an area to identify the Kaiser effect as well as not being able to simultaneously fit for the velocity dispersion.

### 3.5.2 Growth rate results compared

Using the results for  $\beta$  and the galaxy bias we can compare our constraints of the growth parameter  $f(z)$  to previous results. Guzzo et al. (2008) presented the results of such an analysis based on the VLT VIMOS Deep Survey (VVDS), showing values for  $f(z)$  extracted from a number of galaxy surveys up to a redshift of  $z \approx 0.8$ . Here we add the  $z \approx 3$  result from our survey. We present measurements in terms of both  $f(z)$  and  $f\sigma_8$ , where  $f\sigma_8$  is intended to give a measurement which is less dependent on the cosmology assumed for the calculation of the clustering (e.g. Song & Percival 2009).

We have already calculated the infall parameter and take the value ( $\beta = 0.38$ ) obtained via fitting the velocity dispersion and  $\beta$  in a  $\Lambda \text{CDM}$  cosmology with  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$  (Fig. 23). Combining this with our measurement of the galaxy bias gives a value for the growth parameter based on the combined LBG sample of  $f(z=3) = 0.99 \pm 0.50$ .

We present the  $f(z=3)$  result (filled blue circle) in the top panel of Fig. 26 alongside a number of other low-redshift measurements. In order of ascending redshift, the star shows the measurement of Turnbull et al. (2012) based on local supernovae measurements,



**Figure 26.** The evolution of the growth factor based on available  $z < 1$  survey observations and the constraint from the VLRS data (solid blue circle).

the diamond shows the result based on the 2dFGRS presented by Hawkins et al. (2003), the red triangle shows the SDSS result based on LRGs from Tegmark et al. (2006), the red square shows the 2dF-SDSS LRG and QSO (2SLAQ) result also estimated from the LRG population of Ross et al. (2007) and the black circle shows the VADS result of Guzzo et al. (2008). For completeness these are also summarized in Table 12.

In the middle panel of Fig. 26, we also plot the evolution of  $f(z)$  based on the assumed  $\Lambda$ CDM cosmology (dashed line), where  $f(z) = \Omega_m(z)^{0.55}$ . The low-redshift data points are all consistent with the assumed cosmology at the  $\sim 1\sigma$  level and at  $z = 3$ , the model cosmology is again consistent with the data. We note again that the observations themselves depend on the assumed cosmology via

$\sigma_8(z)$  and so to test the  $\Omega_m = 1$  cosmology we adjust the observed values of  $f(z)$  for the effects of different cosmology in equation (14) according to the methods set out by da Ângela et al. (2005b). Also assuming that  $\beta$  is approximately independent of the assumed cosmology, we see that the  $\Omega_m = 1$   $z$ -independent growth rate is apparently rejected by the data. However, if the bias is allowed to float rather than just fit the lowest redshift point then the model may only be rejected at the  $1-2\sigma$  level, consistent with the conclusions from Fig. 25.

If we now consider  $f\sigma_8$ , the observations are now independent of the assumed cosmology, at least given again the assumption that the observed  $\beta$  is approximately cosmology independent. Each of the observational measurements is again plotted in the top panel of Fig. 26, but now in terms of  $f\sigma_8$ . We now plot three test cosmologies for comparison, the  $\Lambda$ CDM used in the top panel (dashed line), plus an Einstein-de-Sitter model ( $\Omega_m = 1$ ,  $\Omega_\Lambda = 0$ , dotted line) and an open Universe without a cosmological constant and a mass density of  $\Omega_m = 0.3$  (dot-dashed line). For each model we incorporate a factor ( $c_{at}$ ) to correct for the cosmology assumed in the measurement of the clustering observations being different from the test cosmology. Each model is thus given by

$$(f\sigma_8)_a = \beta\sigma_{g,a} = \frac{\beta\sigma_{g,t}}{\sqrt{c_{at}}} = \frac{(f\sigma_8)_t}{\sqrt{c_{at}}} = \frac{\Omega_{m,t}(z)^{0.55}\sigma_{g,t}(z)}{\sqrt{c_{at}(z)}}, \quad (16)$$

where an index of ‘t’ denotes a parameter calculated in the test cosmology and an index of ‘a’ denotes a parameter calculated in the assumed cosmology (i.e.  $\Lambda$ CDM). We normalize  $\sigma_8$  to 0.8 at  $z = 0$  and  $\sigma_g$  is effectively  $\sigma_8$  measured for the galaxy population. If we assume a power-law form for the clustering with a slope of  $\gamma = 1.8$ , then following Ballinger, Peacock & Heavens (1996) and da Ângela et al. (2005b)  $c_{at}$  is given by

$$c_{at} = \left( \left( \frac{B_t}{B_a} \right)^2 \frac{A_t}{A_a} \right)^{2/3}, \quad (17)$$

with  $A$  and  $B$  (da Ângela et al. 2005b) given by

$$A = \frac{c}{H_0} \frac{1}{\Omega_\Lambda^0 + \Omega_m^0(1+z)^3}, \quad (18)$$

$$B = \frac{c}{H_0} \int_0^z \frac{dz'}{\Omega_\Lambda^0 + \Omega_m^0(1+z')^3}. \quad (19)$$

We are assuming here that  $\beta$  is independent of cosmology, a reasonable approximation when  $\Omega_m \gtrsim 0.1$ .

**Table 12.** Summary of growth parameter results from the literature.

Survey	$z$	$b$	$\beta$	$f$	$f\sigma_8$
First Amendment SNe <sup>a</sup>	0.025	—	—	—	$0.40 \pm 0.07$
2dFGRS <sup>b</sup>	0.11	$1.15 \pm 0.06$	$0.47 \pm 0.18$	$0.54 \pm 0.21$	$0.41 \pm 0.16$
SDSS LRGs <sup>c</sup>	0.35	$2.13^d$	$0.31 \pm 0.04$	$0.66 \pm 0.07$	$0.44 \pm 0.05$
2SLAQ LRGs <sup>e</sup>	0.55	$2.02 \pm 0.10$	$0.45 \pm 0.05$	$0.91 \pm 0.11$	$0.55 \pm 0.07$
VADS <sup>f</sup>	0.77	$1.30 \pm 0.10$	$0.70 \pm 0.26$	$0.91 \pm 0.07$	$0.50 \pm 0.04$
VLRS+Keck	2.85	$2.59 \pm 0.13$	$0.38 \pm 0.19$	$0.99 \pm 0.50$	$0.26 \pm 0.13$

<sup>a</sup> Turnbull et al. (2012).

<sup>b</sup> Hawkins et al. (2003).

<sup>c</sup> Tegmark et al. (2006).

<sup>d</sup> Song & Percival (2009).

<sup>e</sup> Ross et al. (2007).

<sup>f</sup> Guzzo et al. (2008).



**Table 13.** Results of the halo matching analysis.

Sample	Bias	$\log(\langle M_h \rangle / h^{-1} M_\odot)$	$\log(M_{\min} / h^{-1} M_\odot)$	$\langle N(M) \rangle$
VLRS	$2.37 \pm 0.21$	$11.57 \pm 0.15$	$11.13 \pm 0.18$	$1.2 \pm 0.6$
Keck	$2.78 \pm 0.13$	$11.69 \pm 0.10$	$11.30 \pm 0.10$	$1.8 \pm 0.6$
Combined	$2.59 \pm 0.13$	$11.73 \pm 0.07$	$11.33 \pm 0.09$	$2.0 \pm 0.5$

With the models corrected to account for differences between the cosmology assumed for the observations and the model cosmologies, the observations now provide clearer tests on the models. We note that the observations (excepting the SDSS point) assume the  $\Omega_m = 0.3/\Omega_\Lambda = 0.7$  cosmology and so there is no change in the relationship between the observations and the  $\Lambda$ CDM model between the top and bottom panels. Thus the VLRS data point shows the same level of consistency with the  $\Lambda$ CDM model for  $f\sigma_8$  and  $f(z)$ .

At redshifts of  $z < 1$ , we see that flat and open cosmologies (i.e. with and without a cosmological constant) are poorly distinguished by the available observations. At  $z = 3$ , we find that the VLRS data can only reject the open cosmology with  $\Omega_m = 0.3$  at the  $\approx 1\sigma$  level. The Einstein–de Sitter cosmology is apparently rejected by combining the  $z < 1$  and  $z \approx 3$  observations. But again if the normalization of the model is allowed to float rather than be fixed on the low-redshift SNe observation, the model still fits the data with a reduced  $\chi^2$  of 2.7.

### 3.6 The dark matter haloes of $z \sim 3$ star-forming galaxies

We now look at the nature of the haloes that host the LBG sample based on our clustering results using the halo occupation distribution (HOD) model formalism (e.g. Ma & Fry 2000; Peacock & Smith 2000; Zehavi et al. 2004; Zheng et al. 2005). By matching our clustering results to the measured clustering properties of simulated dark matter haloes, we estimate mean halo masses ( $\langle M_h \rangle$ ), minimum halo masses ( $M_{\min}$ ) and cumulative occupation numbers ( $\langle N(M) \rangle$ ) for the VLRS, Keck and combined samples. The simulation results are obtained from the cosmological dark matter simulation described by González & Padilla (2010) and the results are given in Table 13.

We find that the results are consistent within the error bars between the VLRS and Keck samples evaluated separately, with mean host halo masses of  $\sim 10^{11.6} h^{-1} M_\odot$ . This lends additional credence to our having combined the two samples in order to improve the statistical fidelity of the 2D clustering results. In addition we note that the occupation numbers suggest that multiple LBGs are present in single galaxies, with  $\langle N(M) \rangle$  consistently  $> 1$  for all the samples although this is with relatively large uncertainties.

There are few other measurements of the halo masses of  $z \sim 3$  LBGs available in the literature that are based on spectroscopic data. In terms of photometric samples, Foucaud et al. (2003), Hildebrandt et al. (2007) and Yoshida et al. (2008) measure halo masses of bright  $z \approx 3$  LBG samples of  $M_{\text{DM}} \sim 10^{12} h^{-1} M_\odot$ , an order of magnitude larger than for our sample. However, in a similarly photometric study, Lee et al. (2006) found marginally lower halo masses of  $\sim 5\text{--}10 \times 10^{11} h^{-1} M_\odot$  for both  $z \sim 3$  and  $\sim 4$  LBGs, the  $z \sim 3$  LBGs having a magnitude limit of  $r = 25.5$ . Similarly, the results of Trainor & Steidel (2012) show a halo mass of  $M_{\text{DM}} \sim 10^{11.9 \pm 0.1} h^{-1} M_\odot$ , but is based on galaxies with a redshift distribution somewhat lower than our own.

In terms of the spectroscopic samples closest in redshift and form to our own, work using the Steidel et al. (2003, 2004) data

report halo masses of  $M_{\text{DM}} \sim 10^{11.5 \pm 0.3} h^{-1} M_\odot$  (Adelberger et al. 2005a). These spectroscopic  $z \sim 3$  based measurements are in good agreement with our own results. This is as one would expect for our ‘Keck’ sample given that this uses some of the same data as the above results, whilst the consistency between these results and the result from our own pure-VLRS sample adds weight to the results as a whole. As noted by previous authors, the LBG host halo masses are approximately an order of magnitude lower than those measured for the infrared selected population at  $z \sim 3$  (e.g. Quadri et al. 2007), hinting at the continued trend for a ‘blue’ star-forming population existing in low-density environments and a ‘red’, potentially more passive population inhabiting denser environments.

## 4 CONCLUSIONS

We have presented the widest area spectroscopic survey of galaxies thus far in the redshift range  $2 < z < 3.5$ , based on observations with the VLT VIMOS instrument. This paper adds to the initial data set of Paper I, where data in five  $0.5 \times 0.5$  fields were presented. Here we add a further four new fields, each with deep optical imaging over an area of  $\approx 0.5 \times 0.5$  in three cases and a full  $1^\circ \times 1^\circ$  in the fourth field. In addition, we have extended one of the original fields of Paper I to  $1^\circ \times 1^\circ$  from the original  $0.5 \times 0.5$ . In total therefore, we now have  $\approx 4 \text{ deg}^2$  of optical imaging with a minimum of three bands in each field incorporating *U*, *B* and *R* or equivalents.

In total, the survey now consists of 1994 spectroscopically confirmed  $z > 2$  galaxies. The properties of the full sample have been presented here with redshift and magnitude distributions as well as example and composite spectra. The mean redshift of our  $z > 2$  galaxy data set is  $\bar{z} = 2.79$ . In addition, we detect 30 AGN or quasars,  $\approx 800$  low-redshift galaxies and  $\approx 130$  Galactic stars. Using the  $z > 2$  galaxy data set, we have conducted an analysis of the galaxy clustering at  $z \sim 3$ . Using the semiprojected correlation function, we have measured a galaxy clustering length of  $r_0 = 3.46 \pm 0.41 h^{-1} \text{ Mpc}$  with a slope of  $\gamma = 1.52 \pm 0.26$ , assuming a power-law form to  $\xi(r)$ . We have also combined the VLRS sample with the Keck LBG sample of Steidel et al. (2003), which provides greater statistical power on small scales (i.e.  $s \lesssim 2 h^{-1} \text{ Mpc}$ ) than the VLRS but does not provide the coverage of the VLRS at larger scales (i.e.  $s \gtrsim 8 h^{-1} \text{ Mpc}$ ). For the combined sample we measure a clustering length of  $r_0 = 3.83 \pm 0.24 h^{-1} \text{ Mpc}$ , with a slope of  $\gamma = 1.60 \pm 0.09$ .

We have shown that the LBG correlation functions consistently show slopes that are significantly flatter than the canonical  $\gamma = 1.8$  observed at low redshift. Indeed, the measured slopes of  $\gamma = 1.55$  are flatter than in some  $z \approx 1$  galaxy and radio-source correlation functions that have been interpreted as showing evidence for primordial non-Gaussianity (Xia et al. 2010; Sawangwit et al. 2011; Thomas et al. 2011; Nikoloudakis et al. 2012). Non-Gaussianity is expected to be easier to detect at large scales and high redshift. We have therefore checked whether a standard  $\Lambda$ CDM model is consistent with the form of the VLRS  $\xi(s)$  in particular in the regime  $10 < s < 50 h^{-1} \text{ Mpc}$ . We found that there is evidence that the LBGs are showing more large-scale power than the standard model in this regime but only at  $\approx 2\sigma$ . More studies of LBG clustering at large scales are clearly needed to check these results.

In addition to the 1D clustering analyses, we have also investigated the 2D correlation function and the imprints of galaxy dynamics on the clustering. We find that the 2D clustering for the VLRS+Keck LBG sample is well fit by a model based on a power-law fit with a clustering length of  $r_0 = 3.46 h^{-1} \text{ Mpc}$ , a large-scale

infall parameter of  $\beta = 0.38 \pm 0.19$  and a velocity dispersion of  $\sqrt{\langle w_z^2 \rangle} = 420_{-160}^{+140} \text{ km s}^{-1}$ , over a range of  $1 < s < 25 h^{-1} \text{ Mpc}$ . We have shown that this result is consistent with the model for the redshift space correlation function,  $\xi(s)$ , measured for the combined sample.

We use the 2D galaxy clustering results to determine the matter density parameter and the growth parameter. Using the previously constrained form for the clustering and galaxy velocity dispersion, we fit the 2D correlation function for the matter density,  $\Omega_m$ . We find an acceptable range in the matter density of  $\Omega_m(z=0) = 0.08_{-0.08}^{+0.22}$  (with an infall parameter of  $\beta(z=3) = 0.38_{-0.13}^{+0.15}$ ). We add a further constraint provided by the 2dFGRS low-redshift clustering measurements, which gives  $\Omega_m(z=0) = 0.30_{-0.18}^{+0.32}$  (with an infall parameter of  $\beta(z=3) = 0.38_{-0.09}^{+0.16}$ ). Although the constraints on the mass density are relatively weak, we see that the constraints on the infall parameter remain consistent. Using these measurements to constrain the growth parameter, we find a value of  $f(z=3) = b\beta = 0.99 \pm 0.50$ . In addition we determine the combined parameter  $f\sigma_8$ , which gives a measure of the growth parameter that is less dependent on the assumed underlying dark matter mass distribution. In this case we find a value of  $f\sigma_8 = 0.26 \pm 0.13$ . These measurements are the highest redshift constraint on the growth parameter based on galaxy clustering analyses. We have shown that these measurements are consistent with the  $\Lambda$ CDM standard model, although given the uncertainties on the measurements, they are also consistent with a number of other cosmologies.

Based on the clustering results, we estimate typical halo masses for the dark matter haloes that host the LBG population. For the VLSR sample alone, we estimate a mean halo mass of  $M_{\text{DM}} = 10^{11.57 \pm 0.15}$ , consistent with measurements based on comparable spectroscopic LBG samples at  $z \sim 3$  and an order of magnitude lower than the typical halo masses hosting  $z \sim 3$  infrared-selected galaxies.

This work is one of the largest surveys of the galaxy mass distribution at  $z \approx 3$  and paves the way for a number of lines of research, which will be followed in subsequent papers. In particular, the proximity of the data presented here to quasar sightlines will provide important constraints on the relationship between galaxies and the IGM at an epoch associated with significant interactions between the two.

## ACKNOWLEDGMENTS

We thank Mike Irwin for assistance with the WFCAM data reduction. This work was based on data obtained with the NOAO Mayall 4-m Telescope at Kitt Peak National Observatory, USA (programme ID: 06A-0133), the NOAO Blanco 4-m Telescope at Cerro Tololo Inter-American Observatory, Chile (programme IDs: 03B-0162, 04B-0022) and the ESO VLT, Paranal, Chile (programme IDs: 075.A-0683, 077.A-0612, 079.A-0442). MDH acknowledges the support of a STFC PhD Studentship grant, whilst RB, TS and NM acknowledge STFC funding. RB acknowledges support from a grant obtained from the Agence Nationale de la Recherche (ANR, France). This work was partially supported by the Consejo Nacional de Investigaciones Científicas y Técnicas and Secretaría de Ciencia y Técnica de la Universidad Nacional de Córdoba, and the European Union Alfa II Programme, through LENAC, the Latin American-European Network for Astrophysics and Cosmology. DM and LI are supported by FONDAF CFA 15010003 and BASAL CATA PFB-06. This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion

Laboratory, California Institute of Technology, under contract with NASA.

## REFERENCES

- Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, *ApJ*, 584, 45
- Adelberger K. L., Steidel C. C., Shapley A. E., Hunt M. P., Erb D. K., Reddy N. A., Pettini M., 2004, *ApJ*, 607, 226
- Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005a, *ApJ*, 619, 697
- Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005b, *ApJ*, 629, 636
- Ballinger W. E., Peacock J. A., Heavens A. F., 1996, *MNRAS*, 282, 877
- Baugh C. M., Gardner J. P., Frenk C. S., Sharples R. M., 1996, *MNRAS*, 283, L15
- Bertin E., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, *ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV*. Astron. Soc. Pac., San Francisco, p. 112
- Bertin E., Arnouts S., 1996, *A&AS*, 317, 393
- Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, *ASP Conf. Ser. Vol. 281, Astronomical Data Analysis Software and Systems XI*. Astron. Soc. Pac., San Francisco, p. 228
- Bielby R., Shanks T., Sawangwit U., Croom S. M., Ross N. P., Wake D. A., 2010, *MNRAS*, 403, 1261
- Bielby R. M. et al., 2011, *MNRAS*, 414, 2 (Paper I)
- Blake C. et al., 2009, *MNRAS*, 395, 240
- Bolzonella M., Miralles J., Pelló R., 2000, *A&A*, 363, 476
- Bruzual A. G., Charlot S., 2003, *ApJ*, 405, 538
- Conroy C., Shapley A. E., Tinker J. L., Santos M. R., Lemson G., 2008, *ApJ*, 679, 1192
- Cooke J., Wolfe A. M., Gawiser E., Prochaska J. X., 2006, *ApJ*, 652, 994
- Cowie L. L., Hu E. M., 1998, *AJ*, 115, 1319
- Crain R. A. et al., 2009, *MNRAS*, 399, 1773
- Crichton N. H. M. et al., 2011, *MNRAS*, 414, 28 (Paper II)
- da Ângela J., Outram P. J., Shanks T., Boyle B. J., Croom S. M., Loaring N. S., Miller L., Smith R. J., 2005a, *MNRAS*, 360, 1040
- da Ângela J., Outram P. J., Shanks T., 2005b, *MNRAS*, 361, 879
- Foucaud S., McCracken H. J., Le Fèvre O., Arnouts S., Brodwin M., Lilly S. J., Crampton D., Mellier Y., 2003, *A&A*, 409, 835
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748
- Gawiser E. et al., 2006, *ApJ*, 642, L13
- Gawiser E. et al., 2007, *ApJ*, 671, 278
- Gialavisco M., Dickinson M., 2001, *ApJ*, 550, 177
- González R. E., Padilla N. D., 2010, *MNRAS*, 407, 1449
- González J. E., Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2012, *MNRAS*, 423, 3709
- Groth E. J., Peebles P. J. E., 1977, *ApJ*, 217, 385
- Guzzo L. et al., 2008, *Nat*, 451, 541
- Hainline K. N., Shapley A. E., Greene J. E., Steidel C. C., 2011, *ApJ*, 733, 31
- Hainline K. N., Shapley A. E., Greene J. E., Steidel C. C., Reddy N. A., Erb D. K., 2012, *ApJ*, 760, 74
- Hawkins E. et al., 2003, *MNRAS*, 346, 78
- Hildebrandt H., Pielorz J., Erben T., Schneider P., Eifler T., Simon P., Dietrich J. P., 2007, *A&A*, 462, 865
- Hildebrandt H., Pielorz J., Erben T., van Waerbeke L., Simon P., Capak P., 2009, *A&A*, 498, 725
- Hildebrandt H. et al., 2012, *MNRAS*, 421, 2355
- Hoyle F., Outram P. J., Shanks T., Boyle B. J., Croom S. M., Smith R. J., 2002, *MNRAS*, 332, 311
- Jenkins A. et al., 1998, *ApJ*, 499, 20
- Jose C., Subramanian K., Srianand R., Samui S., 2012, preprint, arXiv:1208.2097

- Kim H., Baugh C. M., Cole S., Frenk C. S., Benson A. J., 2009, *MNRAS*, 400, 1527
- Lee K.-S., Giavalisco M., Gnedin O. Y., Somerville R. S., Ferguson H. C., Dickinson M., Ouchi M., 2006, *ApJ*, 642, 63
- Le Fèvre O. et al., 2003, in Iye M., Moorwood A. F. M., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conf. Ser. Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*. SPIE, Bellingham, p. 1670
- Ma C.-P., Fry J. N., 2000, *ApJ*, 543, 503
- McCracken H. J., Le Fèvre O., Brodwin M., Foucaud S., Lilly S. J., Cramp-ton D., Mellier Y., 2001, *A&A*, 376, 756
- Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, *MNRAS*, 283, 1388
- Massey R. et al., 2007, *Nat*, 445, 286
- Mo H. J., White S. D. M., 1996, *MNRAS*, 282, 347
- Nikoloudakis N., Shanks T., Sawangwit U., 2012, preprint, arXiv:1204.3609
- Norberg P. et al., 2002, *MNRAS*, 332, 827
- Orsi A., Lacey C. G., Baugh C. M., Infante L., 2008, *MNRAS*, 391, 1589
- Ouchi M. et al., 2005, *ApJ*, 635, L117
- Ouchi M. et al., 2008, *ApJS*, 176, 301
- Peacock J. A., Smith R. E., 2000, *MNRAS*, 318, 1144
- Peebles P. J. E., 1980, *The Large-Scale Structure of the Universe*. Princeton Univ. Press, Princeton, NJ
- Phleps S., Peacock J. A., Meisenheimer K., Wolf C., 2006, *A&A*, 457, 145
- Quadri R. et al., 2007, *ApJ*, 654, 138
- Quadri R. F., Williams R. J., Lee K., Franx M., van Dokkum P., Brammer G. B., 2008, *ApJ*, 685, L1
- Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, *ApJS*, 175, 48
- Roche N., Shanks T., Metcalfe N., Fong R., 1993, *MNRAS*, 263, 360
- Ross N. P. et al., 2007, *MNRAS*, 381, 573
- Savoy J., Sawicki M., Thompson D., Sato T., 2011, *ApJ*, 737, 92
- Sawangwit U., Shanks T., Abdalla F. B., Cannon R. D., Croom S. M., Edge A. C., Ross N. P., Wake D. A., 2011, *MNRAS*, 416, 3033
- Shanks T., Bielby R., Infante L., 2011, *Messenger*, 143, 42
- Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, *ApJ*, 588, 65
- Smith R. E. et al., 2003, *MNRAS*, 341, 1311
- Song Y.-S., Percival W. J., 2009, *J. Cosmol. Astropart. Phys.*, 10, 4
- Springel V. et al., 2005, *Nat*, 435, 629
- Springel V., Frenk C. S., White S. D. M., 2006, *Nat*, 440, 1137
- Steidel C. C., Hamilton D., 1993, *AJ*, 105, 2017
- Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, *ApJ*, 462, L17
- Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, *ApJ*, 519, 1
- Steidel C. C., Hunt M. P., Shapley A. E., Adelberger K. L., Pettini M., Dickinson M., Giavalisco M., 2002, *ApJ*, 576, 653
- Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, *ApJ*, 592, 728
- Steidel C. C., Shapley A. E., Pettini M., Adelberger K. L., Erb D. K., Reddy N. A., Hunt M. P., 2004, *ApJ*, 604, 534
- Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289
- Tegmark M. et al., 2006, *Phys. Rev. D*, 74, 123507
- Thomas S. A., Abdalla F. B., Lahav O., 2011, *Phys. Rev. Lett.*, 106, 241301
- Trainor R. F., Steidel C. C., 2012, *ApJ*, 752, 39
- Turnbull S. J., Hudson M. J., Feldman H. A., Hicken M., Kirshner R. P., Watkins R., 2012, *MNRAS*, 420, 447
- Xia J.-Q., Viel M., Baccigalupi C., De Zotti G., Matarrese S., Verde L., 2010, *ApJ*, 717, L17
- Yoshida M., Shimasaku K., Ouchi M., Sekiguchi K., Furusawa H., Okamura S., 2008, *ApJ*, 679, 269
- Zehavi I. et al., 2004, *ApJ*, 608, 16
- Zheng Z. et al., 2005, *ApJ*, 633, 791

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.